



PA (STER-) STERRENBEID BIOTECHNOLOGIE NORTH AMERICA.  
 XX Carcagno CM, Criscuolo M, Melo C, Vidal JA;  
 PI WPI: 2000-376485/32.  
 XX

PT Novel methods for purifying recombinant human erythropoietin from  
 XX mammalian cell culture reagents -  
 XX Claim 16; Page 18; 30pp; English.

CC The present invention relates to a method for purifying erythropoietin  
 CC (EPO) for treatment of disease, especially anaemia. The method involves  
 CC treating cell culture supernatants with differential precipitation,  
 CC hydrophobic interaction chromatography, diafiltration, anionic and  
 CC cationic exchange chromatography and molecular exclusion  
 CC chromatography. The present sequence is the protein from the culture  
 CC supernatant of transfected cell lines, after purification by the above  
 CC process. The sequence shows total homology with natural human EPO.  
 CC The advantage of this method is that high purity and quality EPO is  
 CC produced. A further advantage is that the process does not involve the  
 CC use of organic solvents that may harm the environment.

XX Sequence 165 AA;

Query Match Best Local Similarity 100.0%; Score 846; DB 21; Length 165;  
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVNFFYAMKRMVEGQQA 60  
 DB 1 APPRICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVNFFYAMKRMVEGQQA 60  
 OY 61 VEWVGGLALSEAVLRGQALLVNSQWPEPLQHVDAKAVSGLSLTLRALGAQKEAIS 120  
 DB 61 VEWVGGLALSEAVLRGQALLVNSQWPEPLQHVDAKAVSGLSLTLRALGAQKEAIS 120  
 OY 121 PPDASAAPLRTITADTFRKLFVYSNPLRGKLYTGACRTGD 165  
 DB 121 PPDASAAPLRTITADTFRKLFVYSNPLRGKLYTGACRTGD 165

RESULT 2

ID AAB03760 standard; protein; 165 AA.

XX AAB03760;

DT 04-OCT-2000 (first entry)

DE Human erythropoietin (EPO) amino acid sequence.

KW Erythropoietin; EPO; human; erythroblast differentiation; anaemia;

XX large scale production; renal failure.

OS Homo sapiens.

PN WO200027997-A1.

PD 18-MAY-2000.

PF 08-NOV-1999; 99WO-US26240.

PR 06-NOV-1998; 98AR-0105611.

PR 23-FEB-1999; 99AR-0100683.

PA (STER-) STERRENBEID BIOTECHNOLOGIE NORTH AMERICA.

PI Carcagno CM, Criscuolo M, Melo C, Vidal JA;

DR WPI: 2000-376519/32.

PT A novel method for the massive culture of recombinant mammalian cells

PT producing recombinant human erythropoietin -  
 XX Example 8; Page 11-12; 23pp; English.

CC This sequence represents the human erythropoietin amino acid sequence.  
 CC Erythropoietin is a glycoprotein that stimulates erythroblast  
 CC differentiation in the bone marrow. The present invention relates to a  
 CC method for the large scale production of human EPO from recombinant  
 CC mammalian cells. The method comprises culturing mammalian cells which  
 CC express recombinant human EPO in culture medium comprising insulin.  
 CC Erythropoietin can be used to treat anaemia derived from renal failure.  
 CC The method allows for the industrial scale production of EPO, and  
 CC overcomes the problems of low reproducibility and output quality which  
 CC are encountered with previous production methods.

XX Sequence 165 AA;

Query Match Best Local Similarity 100.0%; Score 846; DB 21; Length 165;  
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLCDRSYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVNFFYAMKRMVEGQQA 60  
 DB 1 APPRLCDRSYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVNFFYAMKRMVEGQQA 60  
 OY 61 VEWVGGLALSEAVLRGQALLVNSQWPEPLQHVDAKAVSGLSLTLRALGAQKEAIS 120  
 DB 61 VEWVGGLALSEAVLRGQALLVNSQWPEPLQHVDAKAVSGLSLTLRALGAQKEAIS 120  
 OY 121 PPDASAAPLRTITADTFRKLFVYSNPLRGKLYTGACRTGD 165  
 DB 121 PPDASAAPLRTITADTFRKLFVYSNPLRGKLYTGACRTGD 165

RESULT 3

ID AAY99705 standard; protein; 165 AA.

XX AAY99705;

DT 15-SEP-2000 (first entry)

DE Non-glycosylated erythropoietin analogue NCB-166delta.

KW Human; non-glycosylated erythropoietin analogue; NGEA; haematocrit;  
 XX antianemic; anaemia; erythropoiesis promoter; mutant; mutein.

OS Homo sapiens.

PN WO200032772-A2.

PD 08-JUN-2000.

PF 23-NOV-1999; 99WO-US27801.

PR 30-NOV-1998; 98US-0110289.

PA (ELIL ) LILLY & CO ELI.

PI Beals JM, Glaesner W, Micranovic R, Milligan RL, Wlitcher DR;

DR WPI: 2000-412320/35.

DR N-PSDB; AAA48373.

PT Non-glycosylated erythropoietic compound useful for increasing  
 PT hematocrit level in mammal with insufficient hematocrit levels in  
 PT conditions such as anemia, comprises protein covalently bonded to  
 PT polymer -

PS Claim 2; Page 93-94; 94pp; English.

CC The present sequence is a non-glycosylated erythropoietin analogue

CC (NGBA) designated NGE-166delta. The protein sequence is identical to  
CC the sequence of wild-type human non-glycosylated erythropoietin NGE  
CC except that Arg at position 166 is deleted. NGE promotes erythropoiesis  
CC and can therefore be used to increase haematocrit levels in mammals  
CC with conditions such as anaemia, in which levels of haematocrit are  
CC insufficient. NGE analogues can also be used to treat such conditions.  
CC NGEAs do not themselves cause a significant increase in haematocrit but  
CC they acquire that property once they are derivatised with polyethylene  
CC glycol polymers. The analogues can be produced using a linkerless  
CC aldehyde modification process. They show stability and bioactivity in  
CC vivo. The nucleotide sequence encoding this protein was constructed  
CC synthetically by in vitro hybridisation using a set of six overlapping  
CC oligonucleotides from the positive strand of human erythropoietin cDNA  
CC with six complementary oligonucleotides (negative strand). The codon  
CC usage was 100% optimised for E. coli codon usage. The hybridised  
CC oligonucleotides were ligated with T4 DNA ligase and the ligation product  
CC amplified by PCR. The nucleotide sequence was used to express the protein  
CC in host cells.

SQ Sequence 165 AA:

Query Match 100.0%; Score 846; DB 21; Length 165;  
Best Local Similarity 100.0%; Pred. No. 1.1e-86;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLCDRSRVLERYLLEKAEKNITTCGAHCSTLNENITVPDTKYNFYAMKRMVEGQQA 60  
DB 1 APPRLCDRSRVLERYLLEKAEKNITTCGAHCSTLNENITVPDTKYNFYAMKRMVEGQQA 60  
QY 61 VEWMOGLALISEAVNLGQALLVNSSQPWEPQLQHVDAKAVSGLSLTTLLRALGAQKEAIS 120  
DB 61 VEWMOGLALISEAVNLGQALLVNSSQPWEPQLQHVDAKAVSGLSLTTLLRALGAQKEAIS 120  
QY 121 PPDASAAPLRTITADTFPRKLFRRVYSNPLRGKIKLYTGECACRTGD 165  
DB 121 PPDASAAPLRTITADTFPRKLFRRVYSNPLRGKIKLYTGECACRTGD 165

RESULT 4  
AAV93445  
ID AAV93445 standard; protein; 165 AA.

AC AAV93445;  
XX  
XX  
DT 04-SEP-2000 (first entry)

XX Amino acid sequence of human erythropoietin.

XX Human; erythropoietin; EPO; anaemia; renal failure.

XX Homo sapiens.

XX WO200028066-A1.

XX 18-MAY-2000.

XX 08-NOV-1999; 99WO-US26238.

XX 06-NOV-1998; 98AR-0105609.

XX 23-FEB-1999; 99AR-0100679.

XX (STER-) STERRENBEID BIOTECHNOLOGIE NORTH AMERICA.

XX Carcagno CM, Criscuolo M, Melo C, Vidal JA;

XX WPI; 2000-376574/32.

XX New host cell producing recombinant human erythropoietin (EPO) used for

XX large scale production of EPO -

XX Claim 1; Page 26-27; 51pp; English.

XX The present sequence represents human erythropoietin protein. The

CC specification describes a host cell line which is used to produce  
CC human erythropoietin (EPO). EPO is a glycoprotein. The cell line  
CC is used for the production of recombinant human erythropoietin. The  
CC protein is used for the treatment of anaemia, especially anaemia derived  
CC from renal failure.

SQ Sequence 165 AA:

Query Match 100.0%; Score 846; DB 21; Length 165;  
Best Local Similarity 100.0%; Pred. No. 1.1e-86;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLCDRSRVLERYLLEKAEKNITTCGAHCSTLNENITVPDTKYNFYAMKRMVEGQQA 60  
DB 1 APPRLCDRSRVLERYLLEKAEKNITTCGAHCSTLNENITVPDTKYNFYAMKRMVEGQQA 60  
QY 61 VEWMOGLALISEAVNLGQALLVNSSQPWEPQLQHVDAKAVSGLSLTTLLRALGAQKEAIS 120  
DB 61 VEWMOGLALISEAVNLGQALLVNSSQPWEPQLQHVDAKAVSGLSLTTLLRALGAQKEAIS 120  
QY 121 PPDASAAPLRTITADTFPRKLFRRVYSNPLRGKIKLYTGECACRTGD 165  
DB 121 PPDASAAPLRTITADTFPRKLFRRVYSNPLRGKIKLYTGECACRTGD 165

RESULT 5  
AAB84525  
ID AAB84525 standard; protein; 165 AA.

XX AAB84525;

XX 05-SEP-2001 (first entry)

XX Amino acid sequence of human erythropoietin (EPO) protein.

XX Erythropoietin; EPO; erythropoietin stimulating protein; NESP;

XX sustained release.

XX Homo sapiens.

XX WO200130320-A1.

XX 03-MAY-2001.

XX 23-OCT-2000; 2000WO-US29257.

XX 22-OCT-1999; 99US-0426566.

XX 13-OCT-2000; 2000US-2222222.

XX (AMGE-) AMGEN INC.

XX Burke P, Klumb L, Murphy K, Herberger J, French DL;

XX WPI; 2001-417552/44.

XX Sustained release composition comprises an active biological

XX ingredient, notably a protein or other biopolymer, particularly

XX erythropoietin stimulating protein, in biocompatible, biodegradable

XX polymeric microparticles -

XX Disclosure; Page 56; 61pp; English.

XX The present sequence encodes a human erythropoietin (EPO) protein.

XX The specification describes a composition for the sustained release of

XX biologically active EPO stimulating protein (NESP). The reduced

XX frequency of administration of NESP which requires preferably injection

XX by skilled personnel, improves patient compliance. Also, sustained

XX release reduces the nature and severity of any side effects of the drug.

XX Sequence 165 AA:

Query Match 100.0%; Score 846; DB 22; Length 165;  
Best Local Similarity 100.0%; Pred. No. 1.1e-86;

|    | Matches | 165,   | conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
|----|---------|--|--------------|----|------------|----|--------|----|------|----|
| QY | 1       | APPRLICRSRYLERTILELKEKEENITTTGCAEHCISLNETITVDTKYNFYAMRMVEYGOOA | 60           |    |            |    |        |    |      |    |
| Db | 1       | APPLICDSRYLERTILELKEKEENITTTGCAEHCISLNETITVDTKYNFYAMRMVEYGOOA  | 60           |    |            |    |        |    |      |    |
| QY | 61      | VEWVGGLALLSAVLRGQALLVNSSQSPWEPDLOHDKRAVSGLSRLTTLRALGAQKEAITS   | 120          |    |            |    |        |    |      |    |
| Db | 61      | VEWVGGLALLSAVLRGQALLVNSSQSPWEPDLOHDKRAVSGLSRLTTLRALGAQKEAITS   | 120          |    |            |    |        |    |      |    |
| QY | 121     | PPDAASAAPLRTITADTFERKLFRVYSNPLRGKLKLYTGACACTGD                 | 165          |    |            |    |        |    |      |    |
| Db | 121     | PPDAASAAPLRTITADTFERKLFRVYSNPLRGKLKLYTGACACTGD                 | 165          |    |            |    |        |    |      |    |

|    |  |
|----|--|
| XX | RESULT 6   |
| XX | AAB66697   |
| XX | ID AAB66697 standard; protein; 165 AA.                                     |
| XX | AC   |
| XX | AD AAB66697;   |
| XX | DT   |
| XX | 06-APR-2001 (first entry)  |
| XX | DE   |
| XX | Human erythropoietin protein #1.   |
| XX | DE   |
| XX | Erythropoietin; EPO; reticulocytes; red blood cell;                        |
| XX | KW ethylene glycol; chronic renal failure; AIDS; cancer.                   |
| XX | OS   |
| XX | Homo sapiens.  |
| XX | PM W0200102017-A2.   |
| XX | 11-JAN-2001.   |
| XX | 28-JUN-2000; 2000WO-EP06009.   |
| XX | 02-JUL-1999; 99US-0142243.   |
| XX | 05-AUG-1999; 99US-0147452.   |
| XX | 30-AUG-1999; 99US-0151454.   |
| XX | (HOFF ) HOFFMANN LA ROCHE & CO AG F.                                       |
| XX | Burg J, Hilger B, Josel H;   |
| XX | WPI: 2001-147051/15.   |
| XX | DR   |
| XX | Novel erythropoietin-glycoprotein conjugate useful for treating            |
| XX | PT diseases correlated with anemia in chronic renal failure patients, AIDS |
| XX | PT and for treating cancer, is linked to polyethylene glycol through       |
| XX | linker -   |
| XX | Claim 19; Fig 1; 40pp; English.  |
| XX | XX   |
| XX | The present invention relates to a conjugate comprising, human             |
| XX | erythropoietin glycoprotein (EPO) having at least one free                 |
| XX | amino group and having in vivo biological activity of causing              |
| XX | an increase the production of reticulocytes and red blood cells,           |
| XX | covalently linked to 1-3 lower-alkoxy poly(ethylene glycol) groups         |
| XX | through a linker. The invention is useful for preparation of               |
| XX | medicaments for the treatment of prophylaxis of disease correlated         |
| XX | with anemia in chronic renal failure patients (CRF), AIDS and for          |
| XX | the treatment of cancer patients undergoing chemotherapy.                  |
| XX | Sequence 165 AA;   |
| XX | 165 AA;  |
| XX | 165 AA;  |

|                       |   |                   |          |            |
|-----------------------|---|-------------------|----------|------------|
| Query Match           | 100.0%  | Score 846         | DB 22    | Length 165 |
| Best Local Similarity | 100.0%  | Pred. NO. 1.1e-86 |          |            |
| Matches 165           | Conservative 0  | Mismatches 0      | Indels 0 | Gaps 0     |
| 1                     | APPRICDSRYLERYLLLEAKAEAMENITTTSCAHCSSINENITVPDTKYNFYAMKRMVEGGOA | 60                |          |            |
| 1                     | APPRICDSRYLERYLLLEAKAEAMENITTTSCAHCSSINENITVPDTKYNFYAMKRMVEGGOA | 60                |          |            |

| QY | 61  | VEWOGALLSEAVLRGCOLLYNSSQPEWPIQLQHYDKAVSGLSLTTLLRNLGAKQKAIS              | 120 |
|----|-----|---|-----|
| QY | 61  | VEWOGALLSEAVLRGCOLLYNSSQPEWPIQLQHYDKAVSGLSLTTLLRNLGAKQKAIS <td>120</td> | 120 |
| Db | 61  | VEWOGALLSEAVLRGCOLLYNSSQPEWPIQLQHYDKAVSGLSLTTLLRNLGAKQKAIS <td>120</td> | 120 |
| QY | 121 | PDDAASAPLRTITADTFRKLFRVYSNFRGKLELTYGCAERTGD                             | 165 |
| Db | 121 | PDDAASAPLRTITADTFRKLFRVYSNFRGKLELTYGCAERTGD                             | 165 |

RESULT 7  
 ABB77896  
 ID ABB77896 standard; protein; 165 AA.  
 AC ABB77896;  
 DT 07-OCT-2002 (first entry)  
 DE Amino acid sequence of a human erythropoietin (EPO).  
 XX  
 XX Human; erythropoietin; EPO; glycoprotein; reticulocyte production;  
 XX red blood cell production; anaemia; chronic renal failure;  
 XX acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;  
 XX committed erythroid progenitor.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200249673-A2.  
 PD 27-JUN-2002.  
 XX  
 PF 08-DEC-2001; 2001WO-EP14434.  
 XX  
 PR 20-DEC-2000; 2000EP-0127891.  
 XX  
 PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
 XX  
 FI Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischner W;  
 FI Wozny M;  
 XX  
 DR WPI: 2002-566640/50.  
 PT  
 PT Novel conjugate of erythropoietin glycoprotein with polyethylene  
 PT glycol, useful for treating diseases correlated with anemia in chronic  
 PT renal failure patients and acquired immunodeficiency syndrome -  
 XX  
 PS Claim 26; Fig 1; 40pp; English.  
 CC  
 CC The present sequence represents a human erythropoietin (EPO) protein.  
 CC It was used to produce conjugates of the invention. The specification  
 CC describes a conjugate comprising an EPO glycoprotein having an N-terminal  
 CC alpha-amino group, chosen from human EPO (hEPO) or its analogues (where  
 CC hEPO is modified by addition of 1-6 glycosylation sites or a  
 CC rearrangement of a glycosylation site). The glycoprotein is covalently  
 CC linked to a poly(ethylene glycol) group. The EPO glycoprotein has in vivo  
 CC biological activity of causing bone marrow cells to increase production  
 CC of reticulocytes and red blood cells. The conjugate increased circulating  
 CC half-life and plasma residence time, decreased clearance, increased  
 CC clinical activity in vivo, improved potency and stability, when compared  
 CC to unmodified EPO. The EPO conjugate is useful for preparing medicaments  
 CC for the treatment and prophylaxis of diseases correlated with anemia in  
 CC chronic renal failure patients (CRF), acquired immunodeficiency syndrome  
 CC (AIDS) and for treating cancer patients undergoing chemotherapy. It is  
 CC also useful for treating patients by stimulating the division and  
 CC differentiation of committed erythroid progenitors in the bone marrow.  
 XX  
 Sequence 165 AA:  
 XX

|                       |                 |   |           |             |
|-----------------------|-----------------|---|-----------|-------------|
| Query Match           | 100.0%;         | Score 846;  | DB 23;    | Length 165; |
| Best Local Similarity | 100.0%;         | Pred. No. 1,1e-86;  |           |             |
| Matches 165;          | Conservative 0; | Mismatches 0;   | Indels 0; | Gaps 0;     |
| QY                    | 1               | APRRICDSRVLEERYLLKEAKENITTTGCAEHCISINENITVPDKVNFYAMKREYGOQA | 60        |             |
|                       |                 |   |           |             |

Db 1 APRRLICDSRYLERYLEAKEAENITTCGAHCSLNENITVPDTKVFYAMKRMEVGQA 60  
 QY 61 VEWOGIALISEAVLRGQALLVNSOPWEPLOLHVDKAVGSLSTTLRALGAQKEAIS 120  
 Db 61 VEWOGIALISEAVLRGQALLVNSOPWEPLOLHVDKAVGSLSTTLRALGAQKEAIS 120  
 QY 121 PPDAASAPLRTITADTFRKLFRYNSNLRGKILKLYTGECRCRGD 165  
 Db 121 PPDAASAPLRTITADTFRKLFRYNSNLRGKILKLYTGECRCRGD 165

RESULT 8  
 AAMS3061  
 ID AAMS3061 standard; protein; 165 AA.  
 AC AAMS3061;  
 XX  
 DT 25-MAR-2002 (first entry)  
 XX  
 DE Human erythropoietin (hepo), 165 residue form.  
 XX  
 KW Human; erythropoietin; EPO; hepo; haemostatic; red blood cell;  
 KW blood disorder; anaemia; chronic renal failure; CRF; AIDS;  
 KM acquired immunodeficiency syndrome; cancer chemotherapy; haemostatic;  
 KM anti-HIV; antianaemic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 7..161  
 FT Modified-site 24  
 FT /note= "N-glycosylated"  
 FT Disulfide-bond 29..33  
 FT Modified-site 38  
 FT /note= "N-glycosylated"  
 FT Modified-site 83  
 FT /note= "N-glycosylated"  
 FT Modified-site 126  
 FT /note= "O-glycosylated"  
 XX  
 PN WO200187329-A1.  
 XX  
 PD 22-NOV-2001.  
 XX  
 PF 08-MAY-2001; 2001WO-EP05187.  
 XX  
 PR 15-MAY-2000; 2000EP-0110355.  
 XX  
 PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
 XX  
 PI Papadimitriou A;  
 XX  
 DR WPI; 2002-082943/11.  
 XX  
 PT Composition useful in the treatment of e.g. AIDS comprises an  
 PT erythropoietin protein, and a multiple charged inorganic anion in a  
 PT buffer -  
 XX  
 PS Claim 28; Fig 1; 64pp; English.  
 XX  
 CC The invention relates to liquid pharmaceutical compositions comprising  
 CC an erythropoietin (Epo) protein, a multiple negatively charged inorganic  
 CC anion in a buffer which maintains the pH of the solution from 5.5-7.0,  
 CC and optionally at least one excipient. The erythropoietin used in the  
 CC composition is preferably human (AAMS3061 or AAMS3062) a human  
 CC erythropoietin variant containing additional glycosylation sites  
 CC (AAMS3064-AAMS3107), or an erythropoietin with the C-terminal addition  
 CC of a C-terminal fragment of human chorionic gonadotropin (AAMS3063).  
 CC Erythropoietin is a glycoprotein essential for the formation of red blood  
 CC cells and is therefore useful in the treatment of blood disorders  
 CC characterised by low or defective red blood cell production. The  
 CC compositions of the invention can be used in the treatment and prevention  
 CC of anaemia in chronic renal failure patients (CRF), AIDS (acquired

CC immunodeficiency syndrome), and/or for the treatment of cancer patients  
 CC undergoing chemotherapy. Unlike prior art erythropoietin compositions,  
 CC the compositions of the invention do not contain human serum albumin  
 CC (thereby avoiding the possibility of viral infections and allergic  
 CC reactions associated with this component), are liquid rather than  
 CC lyophilisates (and therefore do not need to be reconstituted before  
 CC administration), and are stable at elevated temperatures such as 25  
 CC degrees Celsius and even 40 degrees Celsius, and therefore can be stored  
 CC without refrigeration for prolonged periods without degradation and loss  
 CC of activity. The present sequence represents the 165 residue form  
 CC of human erythropoietin which is specifically claimed for use in a  
 CC composition of the invention.  
 XX  
 SQ Sequence 165 AA:  
 QY 100.0%; Score 846; DB 23; Length 165;  
 Best local Similarity 100.0%; Pred. No. 1.1e-86;  
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 APRRLICDSRYLERYLEAKEAENITTCGAHCSLNENITVPDTKVFYAMKRMEVGQA 60  
 Db 1 APRRLICDSRYLERYLEAKEAENITTCGAHCSLNENITVPDTKVFYAMKRMEVGQA 60  
 QY 61 VEWOGIALISEAVLRGQALLVNSOPWEPLOLHVDKAVGSLSTTLRALGAQKEAIS 120  
 Db 61 VEWOGIALISEAVLRGQALLVNSOPWEPLOLHVDKAVGSLSTTLRALGAQKEAIS 120  
 QY 121 PPDAASAPLRTITADTFRKLFRYNSNLRGKILKLYTGECRCRGD 165  
 Db 121 PPDAASAPLRTITADTFRKLFRYNSNLRGKILKLYTGECRCRGD 165

RESULT 9  
 AAP70398  
 ID AAP70398 standard; protein; 166 AA.  
 AC AAP70398;  
 XX  
 DT 19-FEB-1991 (first entry)  
 XX  
 DE Sequence of human erythropoietin (EPO).  
 XX  
 KW Mega-karyocyte-platelet growth factor; hormone;  
 KW mega-karyocyte colony stimulating factor; therapy;  
 KW small acetyl cholinesterase positive cell;  
 KW erythrocyte growth effect.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP62149624-A.  
 XX  
 PD 03-JUL-1987.  
 XX  
 PF 15-AUG-1986; 86JP-0191542.  
 XX  
 PR 13-SEP-1985; 85JP-0203049.  
 XX  
 PR 15-AUG-1986; 86JP-0191542.  
 XX  
 PA (KAWA/) KAWAKITA M.  
 XX  
 DR WPI; 1987-224837/32.  
 XX  
 PT Megakaryocyte-platelet growth factor - contains as active  
 PT component human erythropoietin and is used to treat diseases  
 PT caused by decrease in platelets  
 XX  
 PS Disclosure; Page 181; 8pp; Japanese.  
 XX  
 CC All of the Cys residues in the SQ are labelled "SH". Megakaryocyte-  
 CC platelet growth factor contains human EPO as an active principle.  
 CC Human EPO has a megakaryocyte colony-stimulating activity and  
 CC increases the ratio of small acetyl cholinesterase positive cell  
 CC (SACHe+) which is immature megakaryocyte. Human EPO effects

CC megakaryocyte-platelet system other than an erythrocyte growth  
 CC effect. Megakaryocyte-platelet growth is usable as a remedy for  
 CC diseases caused by a platelet decrease.  
 XX

Sequence 166 AA;

Query Match 100.0%; Score 846; DB 8; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-86;  
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRRLICDSRYLERYLLLEAKAEANITTCGAHCGLNENITVPDTKYVFYAKRMREYGOQA 60  
 |||  
 DB 1 APRRLICDSRYLERYLLLEAKAEANITTCGAHCGLNENITVPDTKYVFYAKRMREYGOQA 60  
 |||  
 QY 61 VEWOGIALISEAVLRGQALLVNSQPWEPLQHLVDKAVSGLSRLTTLRALGAQKEAIS 120  
 |||  
 DB 61 VEWOGIALISEAVLRGQALLVNSQPWEPLQHLVDKAVSGLSRLTTLRALGAQKEAIS 120  
 |||  
 QY 121 PPDAASAAPLRTITADTFPRKLFRRVYSNPLRGKIKLYTGACRTGD 165  
 |||  
 DB 121 PPDAASAAPLRTITADTFPRKLFRRVYSNPLRGKIKLYTGACRTGD 165  
 |||

RESULT 10  
 AAR3593  
 ID AAR3593 standard; Protein; 166 AA.

AC AAR3593;  
 XX  
 DT 20-OCT-1992 (first entry)  
 XX  
 DE Recombinant hematopoietic molecule portion 2.  
 XX  
 KW Erythropoietin; EPO; erythrocytes; IL-3; haematopoiesis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9206116-A.  
 XX  
 PD 16-APR-1992.  
 XX  
 PF 26-SEP-1991; 91WO-US07053.  
 XX  
 PR 28-SEP-1990; 90US-0589958.  
 XX  
 PA (ORTH ) ORTHO PHARM CORP.  
 XX  
 PI Rosen JT;  
 XX  
 DR WPI; 1992-150819/18.  
 XX  
 PT Recombinant haematopoietic molecules useful in treating  
 XX anaemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-5 or M-CSF  
 XX and has early and later myeloid differentiation activity  
 XX  
 PS Disclosure; Page 32; 82pp; English.

CC This protein sequence given comprises the entire amino acid sequence  
 CC of human erythropoietin (EPO). EPO leads to the maturation of  
 CC erythrocytes and is therefore designated as a late myeloid  
 CC differentiation factor (MDF). Within the scope of the invention  
 CC hybrid molecules were produced which contain at least a portion of an  
 CC early MDF and at least a portion of a late MDF covalently linked. The  
 CC EPO sequence given is effective within the scope of the invention in  
 CC full or in a truncated version. Amino acids 7-161 act as a  
 CC late MDF when recombined with an early MDF eg. IL-3.  
 CC These compounds can be used to promote hematopoiesis in a patient.  
 CC The bonding of the early and late factors allows a very high conc. of  
 CC late MDF at the surface of a cell which the early MDF is bound. It  
 CC also allows the early MDA to act more specifically to stimulate only  
 CC the desired lineage, thus reducing undesirable effects. These  
 CC compounds are useful for treating anaemias of various origins eg. renal  
 CC failure and AIDS. It is easier to produce and administer one

CC recombinant molecule rather than two separate molecules.

Sequence 166 AA;

Query Match 100.0%; Score 846; DB 13; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-86;  
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APRRLICDSRYLERYLLLEAKAEANITTCGAHCGLNENITVPDTKYVFYAKRMREYGOQA 60  
 |||  
 DB 1 APRRLICDSRYLERYLLLEAKAEANITTCGAHCGLNENITVPDTKYVFYAKRMREYGOQA 60  
 |||  
 QY 61 VEWOGIALISEAVLRGQALLVNSQPWEPLQHLVDKAVSGLSRLTTLRALGAQKEAIS 120  
 |||  
 DB 61 VEWOGIALISEAVLRGQALLVNSQPWEPLQHLVDKAVSGLSRLTTLRALGAQKEAIS 120  
 |||  
 QY 121 PPDAASAAPLRTITADTFPRKLFRRVYSNPLRGKIKLYTGACRTGD 165  
 |||  
 DB 121 PPDAASAAPLRTITADTFPRKLFRRVYSNPLRGKIKLYTGACRTGD 165  
 |||

RESULT 11  
 AAW7780  
 ID AAW7780 standard; Protein; 166 AA.

AC AAW7780;  
 XX  
 DT 24-NOV-1998 (first entry)  
 XX  
 DE Human EPO receptor agonist polypeptide.  
 XX  
 KW Haematopoietic receptor agonist; erythropoietin receptor agonist;  
 KW EPO; human; chimeric protein; stem cell expansion; tumour;  
 KW infection; autoimmune disease; haematopoietic disorder; therapy;  
 KW dendritic cell.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH Misc-difference 1..6 "1-6 amino acids of the N-terminus are  
 FT optionally deleted"  
 FT Misc-difference 162..166  
 FT "1-5 amino acids of the C-terminus are  
 FT optionally deleted"  
 FT Misc-difference 23..24  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 24..25  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 25..26  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 26..27  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 27..28  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 28..29  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 29..30  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 30..31  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 31..32  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 32..33  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 33..34  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 34..35  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 35..36  
 FT "possible positions of new C- and N-termini"  
 FT Misc-difference 36..37  
 FT "possible positions of new C- and N-termini"

|    |  |               |  |
|----|--|---------------|--|
| FT | Misc-difference                                | 37..38        | "possible positions of new C- and N-terminals" |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 38..39        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 39..40        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 40..41        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 41..42        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 42..43        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 43..44        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 44..45        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 45..46        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 46..47        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 47..48        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 48..49        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 49..50        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 50..51        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 51..52        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 52..53        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 53..54        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 54..55        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 55..56        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 56..57        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 57..58        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 77..78        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 78..79        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 79..80        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 81..82        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 82..83        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 84..85        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 85..86        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 86..87        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 87..88        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 88..89        | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 108..109      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 109..110      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 110..111      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 111..112      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 112..113      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 113..114      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 114..115      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 115..116      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 116..117      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 117..118      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 118..119      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 119..120      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 120..121      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 121..122      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 122..123      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 123..124      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 124..125      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 125..126      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 126..127      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 127..128      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 128..129      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 129..130      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 130..131      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| FT | Misc-difference                                | 131..132      | "possible positions of new C- and N-terminal"  |
| FT | /note-   |               |  |
| PN | WO9817810-A2.                                  |               |  |
| XX |  |               |  |
| PD | 30-Apr-1998.                                   |               |  |
| XX |  |               |  |
| Pf | 23-Oct-1997;                                   | 97MO-US20037. |  |
| XX |  |               |  |
| PR | 25-Oct-1996;                                   | 96US-0029629. |  |
| XX |  |               |  |
| PA | (SEAR ) SEARLE & CO G D.                       |               |  |
| PI | Feng Y, Mckearn JP, McWhirter CA, Minnerly JC, |               |  |
| PL | Staten NR, Streeter PR, Summers NL, Woulfe SL; |               |  |
| XX |  |               |  |
| DR | WPI: 1998-261504/23.                           |               |  |

CC a CSF, a cytokine, a lymphokine, an interleukin and a haematopoietic  
CC growth factor provided that at least R1 or R2 is selected from (a),  
CC (b) or (c) as above. The multi-functional chimeric haematopoietic  
CC receptor agonist can be used to stimulate the production of

Query Match  
Best Local Similarity 100.0%; Score 846; DB 19; Length 166;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLEAKAEANITTCGAHCSLNENITVPDTKYNFYAMKREYVGQA 60  
Db 1 APPRLICDSRVLEERYLLEAKAEANITTCGAHCSLNENITVPDTKYNFYAMKREYVGQA 60

OY 61 VEWOGIALLSFAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTLLRALGAOKEAIS 120  
Db 61 VEWOGIALLSFAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTLLRALGAOKEAIS 120

OY 121 PPDASAAPLRTITADTFRRKLFRVYSNPLRGKRLKLTGTGECRTGD 165  
Db 121 PPDASAAPLRTITADTFRRKLFRVYSNPLRGKRLKLTGTGECRTGD 165

RESULT 12  
AAM58404  
ID AAM58404 standard; Protein; 166 AA.  
XX AAM58404;  
AC  
XX  
DT 12-OCT-1998 (first entry)  
XX  
DE Human erythropoietin.  
XX  
KW Erythropoietin receptor agonist; EPO; human; anaemia;  
KW haematopoietic deficiency; red blood cell; erythroid progenitor;  
KW bone marrow suppression.  
XX  
OS Homo sapiens.  
XX  
PN WO9818926-A1.  
PD 07-MAY-1998.  
XX  
PF 23-OCT-1997; 97WO-US18703.  
XX  
PR 25-OCT-1996; 96US-0034044.  
PA (SEAR ) SEARLE & CO G D.  
XX  
PI Feng Y, McWhorter CA, Summers N;  
XX WPI: 1998-272221/24.  
XX N-PSDB; AAV31031.  
XX  
PT Human erythropoietin receptor agonist polypeptide - used to  
XX stimulate the production of red blood cells in a patient  
XX  
XX Claim 1; Page 93; 112pp; English.

A claimed human erythropoietin (EPO) receptor agonist polypeptide  
CC comprises a modified EPO amino acid sequence given in AAM58404,  
CC where (a) optionally 1-6 amino acids from the N-terminus and 1-5  
CC from the C-terminus can be deleted, (b) the N-terminus is joined to  
CC the C-terminus directly or through a linker (see AAM58405-12) capable  
CC of joining the N-terminus to the C-terminus, (c) there are new C-  
CC 23-2N termini at any two consecutive amino acids from amino acids  
CC 77-78 to 82-83, 84-85 to 86-85, and 108-109 to 131-132, and (d)  
CC optionally the agonist polypeptide is preceded by Met, Ala, or  
CC Metionine, these circularly permuted EPO receptor agonists  
CC (see AAM58413-72) are claimed. Also claimed are: nucleic acid  
CC molecules (see AAV30971-V31030) encoding novel EPO receptor agonists;  
CC a method of producing an EPO receptor agonist using transformed or  
CC transfected host cells; and methods for stimulating the production

CC of haematopoietic cells, for selective ex vivo expansion of  
CC erythroid progenitors, and treating patients having a haematopoietic  
CC disorder using the EPO receptor agonists. The EPO receptor  
CC agonists retain one or more activities of native EPO and may also  
CC show improved haematopoietic cell-stimulating activity and/or an  
CC improved activity profile which may include reduction of undesirable  
CC biological activities associated with native EPO and/or have  
CC improved physical properties such as increased solubility,  
CC stability and refold efficiency.

Sequence 166 AA:  
SQ

Query Match  
Best Local Similarity 100.0%; Score 846; DB 19; Length 166;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLEAKAEANITTCGAHCSLNENITVPDTKYNFYAMKREYVGQA 60  
Db 1 APPRLICDSRVLEERYLLEAKAEANITTCGAHCSLNENITVPDTKYNFYAMKREYVGQA 60

OY 61 VEWOGIALLSFAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTLLRALGAOKEAIS 120  
Db 61 VEWOGIALLSFAVLRGQALLVNSSQPWEPIQLHVDKAVSGLSLTLLRALGAOKEAIS 120

OY 121 PPDASAAPLRTITADTFRRKLFRVYSNPLRGKRLKLTGTGECRTGD 165  
Db 121 PPDASAAPLRTITADTFRRKLFRVYSNPLRGKRLKLTGTGECRTGD 165

RESULT 13  
ABB07030  
ID ABB07030 standard; Protein; 166 AA.  
XX ABB07030;  
AC  
XX  
DT 21-JUN-2002 (first entry)  
XX  
DE Modified erythropoietin related gene protein sequence.  
XX  
KW Modified erythropoietin; EPO.  
XX  
OS Unidentified.  
XX  
PN KR145802-B1.  
XX  
PD 01-AUG-1998.  
XX  
PF 31-MAY-1994; 94KR-0012082.  
XX  
PR 31-MAY-1994; 94KR-0012082.  
XX  
PA (GLDS ) LG CHEM CO LTD.  
XX  
PI Kim C, Song Y, Lee T;  
XX  
XX WPI: 2000-234250/20.  
XX N-PSDB; ABL50878.  
XX  
PT MODIFIED ERYTHROPOIETIN GENE AND EXPRESSION VECTORS THEREOF -  
XX  
XX Disclosure; Page 14; 15pp; Korean.  
XX  
XX The present invention describes modified erythropoietin (EPO) genes  
CC and expression vectors comprising the genes. The present sequence  
CC represents a protein sequence from the present invention.  
XX  
XX Sequence 166 AA;  
SQ

Query Match  
Best Local Similarity 100.0%; Score 846; DB 21; Length 166;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLEAKAEANITTCGAHCSLNENITVPDTKYNFYAMKREYVGQA 60



```

Db      1 APPRLICDSRVLEKRYLLEKAEENITTCGAHCSLNENITVPDTKYNFAMKREYGOQA 60
      |||
QY      61 VEWOGIALLSEAVLNGQALLVNSSQPWEPLQLVHDKAVSGLSLTLLRALGAQKEAIS 120
      |||
Db      61 VEWOGIALLSEAVLNGQALLVNSSQPWEPLQLVHDKAVSGLSLTLLRALGAQKEAIS 120
QY      121 PPDAASAAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRTGD 165
      |||
      121 PPDAASAAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRTGD 165
      |||

RESULT 14
AAE02641
ID      AAE02641 standard; Protein; 166 AA.
XX
XX      AAE02641;
XX
XX      06-AUG-2001 (first entry)
XX
DE      Human erythropoietin (EPO) mature protein.
XX
KW      Human; erythropoietin; EPO; antihaemic; nephroretrophic; anti-HIV;
KW      vaccine; haemostatic; immunoglobulin; Ig; EPO deficient disease;
KW      anaemia; renal failure; Human Immunodeficiency Virus; HIV;
KW      haematopoietic growth factor.
XX
XX      Homo sapiens.
XX
XX      OS
XX      PN      WO200136489-A2.
XX      PD      25-MAY-2001.
XX      PF      03-NOV-2000; 2000WO-EP10843.
XX      PR      12-NOV-1999; 99US-0164855.
XX      PA      (MERE ) MERCK PATENT GMBH.
XX      PI      Hartmann A, Brandt S, Rieke E, Sobel C, Lo K, Way JC, Gillies S;
XX      DR      WPI; 2001-367563/38.
XX      DR      N-PSDB; AAD06893.
XX
PT      Novel modified erythropoietin forms such as fusion proteins, comprising
PT      the portion of an immunoglobulin molecule and a target molecule having
PT      the biological activity of erythropoietin forms -
XX
XX      Example 1; Page 22; 58pp; English.
XX
CC      The present sequence is human erythropoietin (EPO) mature protein. EPO
CC      has improved biological activity and an extended serum half life greater
CC      than 20 hours. The present invention relates to modified EPO forms such
CC      as fusion proteins comprising a FC portion of an immunoglobulin (Ig)
CC      molecule and an EPO molecule (Fc-EPO). The FC portion is fused covalently
CC      through its C-terminus directly or indirectly to the EPO molecule, and
CC      where the FC portion as well as EPO portion may be modified or mutated.
CC      The invention also relates to non-fused EPO molecules which have a
CC      pattern of cysteines or disulphide bonding which is distinct from human
CC      or animal EPO. Pharmaceutical compositions containing EPO are useful in
CC      the treatment of EPO deficient diseases such as anaemia, renal failure,
CC      HIV infection, blood loss and chronic disease that can be treated with
CC      haematopoietic growth factor.
XX
SQ      Sequence 166 AA;
Query Match      100.0%; Score 846; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.le-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 APPRLICDSRVLEKRYLLEKAEENITTCGAHCSLNENITVPDTKYNFAMKREYGOQA 60
      |||
Db      1 APPRLICDSRVLEKRYLLEKAEENITTCGAHCSLNENITVPDTKYNFAMKREYGOQA 60

```

```

QY      61 VEWOGIALLSEAVLNGQALLVNSSQPWEPLQLVHDKAVSGLSLTLLRALGAQKEAIS 120
      |||
Db      61 VEWOGIALLSEAVLNGQALLVNSSQPWEPLQLVHDKAVSGLSLTLLRALGAQKEAIS 120
QY      121 PPDAASAAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRTGD 165
      |||
      121 PPDAASAAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRTGD 165
      |||

RESULT 15
AAB66698
ID      AAB66698 standard; protein; 166 AA.
XX
XX      AAB66698;
XX
XX      06-APR-2001 (first entry)
XX
DE      Human erythropoietin protein #2.
XX
KW      Erythropoietin; EPO; reticulocytes; red blood cell;
KW      ethylene glycol; chronic renal failure; AIDS; cancer.
XX
XX      Homo sapiens.
XX
XX      OS
XX      PN      WO200102017-A2.
XX      PD      11-JAN-2001.
XX      PF      28-JUN-2000; 2000WO-EP06009.
XX      PR      02-JUL-1999; 99US-0142243.
XX      PR      05-AUG-1999; 99US-0147452.
XX      PR      30-AUG-1999; 99US-0151454.
XX
PA      (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
PI      Burg J, Hilger B, Josel H;
XX
XX      WPI; 2001-147051/15.
XX
XX      Novel erythropoietin-glycoprotein conjugate useful for treating
XX      diseases correlated with anemia in chronic renal failure patients, AIDS
XX      and for treating cancer, is linked to polyethylene glycol through
XX      linker -
XX
XX      Claim 19; Fig 2; 40pp; English.
XX
CC      The present invention relates to a conjugate comprising, human
CC      erythropoietin glycoprotein (EPO) having at least one free
CC      amino group and having in vivo biological activity of causing
CC      an increase the production of reticulocytes and red blood cells,
CC      covalently linked to 1-3 lower-alkoxy poly(ethylene glycol) groups
CC      through a linker. The invention is useful for preparation of
CC      medicaments for the treatment of prophylaxis of disease correlated
CC      with anemia in chronic renal failure patients (CRF), AIDS and for
CC      the treatment of cancer patients undergoing chemotherapy.
XX
SQ      Sequence 166 AA;
Query Match      100.0%; Score 846; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.le-86;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 APPRLICDSRVLEKRYLLEKAEENITTCGAHCSLNENITVPDTKYNFAMKREYGOQA 60
      |||
Db      1 APPRLICDSRVLEKRYLLEKAEENITTCGAHCSLNENITVPDTKYNFAMKREYGOQA 60
QY      61 VEWOGIALLSEAVLNGQALLVNSSQPWEPLQLVHDKAVSGLSLTLLRALGAQKEAIS 120
      |||
Db      61 VEWOGIALLSEAVLNGQALLVNSSQPWEPLQLVHDKAVSGLSLTLLRALGAQKEAIS 120
QY      121 PPDAASAAPLRTTTADTFKRLFRVYSNFLRGKILKLTGTGACRTGD 165
      |||

```

Fri Jan 3 10:09:17 2003

us-09-853-731-1.rag

Page 10

Db 121 PDDASAPLRTITADPFKKLFRYNSNLRGKLYTGACRTGD 165

Search completed: January 2, 2003, 15:11:27  
Job time : 32.9033 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:10:20 ; Search time 8.47432 Seconds  
(without alignments)  
368.997 Million cell updates/sec

Title: US-09-853-731-1

Perfect score: 846  
Sequence: 1 APPRLICDSRYLERYLLERLPAK.....SNFLGKLIKLTGECACRTGD 165

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 117078 segs, 18951520 residues

Total number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCCTUS\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                              |
|------------|-------|-------------|--------|-------|--|
| 1          | 846   | 100.0       | 165    | 10    | US-09-853-731-1<br>Sequence 1, Appl1     |
| 2          | 846   | 100.0       | 165    | 12    | US-10-014-363-1<br>Sequence 1, Appl1     |
| 3          | 846   | 100.0       | 166    | 10    | US-09-853-731-2<br>Sequence 2, Appl1     |
| 4          | 846   | 100.0       | 166    | 12    | US-10-014-363-2<br>Sequence 2, Appl1     |
| 5          | 846   | 100.0       | 169    | 12    | US-10-014-363-4<br>Sequence 4, Appl1     |
| 6          | 846   | 100.0       | 174    | 12    | US-10-014-363-3<br>Sequence 3, Appl1     |
| 7          | 846   | 100.0       | 174    | 12    | US-10-014-363-5<br>Sequence 5, Appl1     |
| 8          | 294   | 34.8        | 60     | 10    | US-09-864-761-48502<br>Sequence 48502, A |
| 9          | 111   | 13.1        | 30     | 10    | US-09-975-063-2<br>Sequence 2, Appl1     |
| 10         | 111   | 13.1        | 30     | 12    | US-10-011-858-2<br>Sequence 2, Appl1     |
| 11         | 103   | 12.2        | 20     | 9     | US-09-919-703-9<br>Sequence 9, Appl1     |
| 12         | 88    | 10.4        | 20     | 9     | US-09-919-703-10<br>Sequence 10, Appl1   |
| 13         | 74    | 8.7         | 1564   | 10    | US-09-801-368-244<br>Sequence 244, App   |
| 14         | 72.5  | 8.6         | 81     | 10    | US-09-764-877-1947<br>Sequence 1947, App |
| 15         | 72    | 8.5         | 321    | 9     | US-09-945-182-26<br>Sequence 26, Appl1   |
| 16         | 70.5  | 8.3         | 813    | 9     | US-09-964-899-25<br>Sequence 25, Appl1   |
| 17         | 70    | 8.3         | 409    | 10    | US-09-815-242-14050<br>Sequence 14050, A |
| 18         | 69.5  | 8.2         | 542    | 9     | US-10-045-815-8<br>Sequence 8, Appl1     |
| 19         | 69.5  | 8.2         | 747    | 9     | US-10-045-815-6<br>Sequence 6, Appl1     |

|    |      |     |      |    |  |
|----|------|-----|------|----|--|
| 20 | 69.5 | 8.2 | 1327 | 10 | US-09-841-835-2<br>Sequence 2, Appl1     |
| 21 | 69.5 | 8.2 | 2472 | 10 | US-09-815-242-5064<br>Sequence 5064, App |
| 22 | 68.5 | 8.1 | 327  | 9  | US-09-916-494A-16<br>Sequence 16, Appl1  |
| 23 | 68   | 8.0 | 2201 | 12 | US-10-029-907-3<br>Sequence 3, Appl1     |
| 24 | 67.5 | 8.0 | 236  | 10 | US-09-905-129-12<br>Sequence 12, Appl1   |
| 25 | 67.5 | 8.0 | 236  | 10 | US-09-905-129-15<br>Sequence 15, Appl1   |
| 26 | 67.5 | 8.0 | 236  | 10 | US-09-991-630-12<br>Sequence 12, Appl1   |
| 27 | 67.5 | 8.0 | 236  | 10 | US-09-991-630-15<br>Sequence 15, Appl1   |
| 28 | 67.5 | 8.0 | 379  | 9  | US-09-975-139-8<br>Sequence 8, Appl1     |
| 29 | 67.5 | 8.0 | 766  | 10 | US-09-925-301-1276<br>Sequence 1276, App |
| 30 | 67   | 7.9 | 136  | 10 | US-09-764-877-1623<br>Sequence 1623, App |
| 31 | 67   | 7.9 | 1148 | 10 | US-09-815-242-13826<br>Sequence 13826, A |
| 32 | 66.5 | 7.9 | 119  | 9  | US-09-738-626-6842<br>Sequence 6842, App |
| 33 | 66.5 | 7.9 | 210  | 10 | US-09-731-872-387<br>Sequence 387, App   |
| 34 | 66   | 7.8 | 319  | 10 | US-09-815-242-10267<br>Sequence 10267, A |
| 35 | 66   | 7.8 | 1618 | 9  | US-09-963-875-1<br>Sequence 1, Appl1     |
| 36 | 66   | 7.8 | 7257 | 9  | US-10-014-717-5<br>Sequence 5, Appl1     |
| 37 | 65.5 | 7.7 | 691  | 10 | US-09-815-242-12339<br>Sequence 12339, A |
| 38 | 65.5 | 7.7 | 711  | 9  | US-09-738-626-3507<br>Sequence 3507, App |
| 39 | 65   | 7.7 | 336  | 10 | US-09-925-300-1408<br>Sequence 1408, App |
| 40 | 65   | 7.7 | 339  | 9  | US-09-981-876-246<br>Sequence 246, App   |
| 41 | 64   | 7.6 | 954  | 9  | US-09-944-413-7<br>Sequence 7, Appl1     |
| 42 | 64   | 7.6 | 954  | 9  | US-09-944-403-7<br>Sequence 7, Appl1     |
| 43 | 64   | 7.6 | 954  | 9  | US-09-944-896-7<br>Sequence 7, Appl1     |
| 44 | 64   | 7.6 | 954  | 9  | US-09-944-944-7<br>Sequence 7, Appl1     |
| 45 | 64   | 7.6 | 954  | 9  | US-09-944-907-7<br>Sequence 7, Appl1     |

#### ALIGNMENTS

```
RESULT 1
US-09-853-731-1
; Sequence 1, Application US/09853731
; Patent No. US20020037841A1
; GENERAL INFORMATION:
; APPLICANT: Papadimitriou, Apollon
; TITLE OF INVENTION: Erythropoietin Composition
; FILE REFERENCE: 20619 US
; CURRENT APPLICATION NUMBER: US/09/853,731
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: EP/00110355.5
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-731-1

Query Match      100.0%; Score 846; DB 10; Length 165;
Best local Similarity 100.0%; Pred. No. 2.7e-84;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 APPRLICDSRYLERYLLERLPAK...
DB      1 APPRLICDSRYLERYLLERLPAK...
QY      61 VEWMOGIALISEAVLNGQALLVNSQPWEPLQIHDVKAAGSLSTTLRALAKQKRAIS 120
DB      61 VEWMOGIALISEAVLNGQALLVNSQPWEPLQIHDVKAAGSLSTTLRALAKQKRAIS 120
QY      121 PPDAASAAPLRTITADTFKLFVYVSNFLGKLIKLTGECACRTGD 165
DB      121 PPDAASAAPLRTITADTFKLFVYVSNFLGKLIKLTGECACRTGD 165
QY      121 PPDAASAAPLRTITADTFKLFVYVSNFLGKLIKLTGECACRTGD 165
DB      121 PPDAASAAPLRTITADTFKLFVYVSNFLGKLIKLTGECACRTGD 165

RESULT 2
US-10-014-363-1
; Sequence 1, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-014-363-1

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 165;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCISLNEITVPDTRKVFYAKRMVEVGOQA 60
    |||||||
DB 1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCISLNEITVPDTRKVFYAKRMVEVGOQA 60
QY 61 VEWOGIALISEAVLRGQALLVNSQWPEPDLQHDKAVSGRLSTTLRLAQAQKEAIS 120
    |||||||
DB 61 VEWOGIALISEAVLRGQALLVNSQWPEPDLQHDKAVSGRLSTTLRLAQAQKEAIS 120
QY 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKILKLTGTGACRTGD 165
    |||||||
DB 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKILKLTGTGACRTGD 165
```

```

RESULT 3
US-09-853-731-2
; Sequence 2, Application US/09853731
; Patent No. US20020037841A1
; GENERAL INFORMATION:
; APPLICANT: Papadimitriou, Apollon
; TITLE OF INVENTION: Erythropoietin Composition
; FILE REFERENCE: 20619 US
; CURRENT APPLICATION NUMBER: US/09/853,731
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: EP/00110355.5
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 166
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-853-731-2
```

```

Query Match
Best Local Similarity 100.0%; Score 846; DB 10; Length 166;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCISLNEITVPDTRKVFYAKRMVEVGOQA 60
    |||||||
DB 1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCISLNEITVPDTRKVFYAKRMVEVGOQA 60
QY 61 VEWOGIALISEAVLRGQALLVNSQWPEPDLQHDKAVSGRLSTTLRLAQAQKEAIS 120
    |||||||
DB 61 VEWOGIALISEAVLRGQALLVNSQWPEPDLQHDKAVSGRLSTTLRLAQAQKEAIS 120
QY 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKILKLTGTGACRTGD 165
    |||||||
DB 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKILKLTGTGACRTGD 165
```

```

RESULT 4
US-10-014-363-2
; Sequence 2, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 166
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-014-363-2

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 166;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCISLNEITVPDTRKVFYAKRMVEVGOQA 60
    |||||||
DB 1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCISLNEITVPDTRKVFYAKRMVEVGOQA 60
QY 61 VEWOGIALISEAVLRGQALLVNSQWPEPDLQHDKAVSGRLSTTLRLAQAQKEAIS 120
    |||||||
DB 61 VEWOGIALISEAVLRGQALLVNSQWPEPDLQHDKAVSGRLSTTLRLAQAQKEAIS 120
QY 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKILKLTGTGACRTGD 165
    |||||||
DB 121 PPDAASAAPLRTITADTFRKLFRVYSNPLRGKILKLTGTGACRTGD 165
```

```

RESULT 5
US-10-014-363-4
; Sequence 4, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 169
; TYPE: PRF
; ORGANISM: CHO/dhfr-
US-10-014-363-4
```

```

Query Match
Best Local Similarity 100.0%; Score 846; DB 12; Length 169;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCISLNEITVPDTRKVFYAKRMVEVGOQA 60
    |||||||
DB 4 APPRLICDSRVLEERYLLEAKAEENITTCGAHCISLNEITVPDTRKVFYAKRMVEVGOQA 63
```

|    |     |  |     |
|----|-----|--|-----|
| OY | 61  | VFWOGJALISEANLRGOCALLVNSQGWPELOI.HVKAASGLSSTITLIRALGAOKEATS  | 120 |
| Dd | 64  | VFWOGJALISEANLRGOCALLVNSQGWPELOIHVKAKVASGLSSTITLIRALGAOKEAIS | 123 |
| OY | 121 | PPDAA\$AAPLRTITADPFRRKLFERYYSNFKGLTKLYTGEACRIGD              | 165 |
| Dd | 124 | PPDAA\$AAPLRTITADPFRRKLFERYYSNFKGLTKLYTGEACRIGD              | 168 |

```

RESULT 6
US-10-014-363-3
: Sequence 3, Application US/10014363
: Patent No. US20020115833A1
: GENERAL INFORMATION:
: APPLICANT: Burg, Josef
: APPLICANT: Engel, Alfred
: APPLICANT: Franze, Reinhard
: APPLICANT: Hilger, Bernd
: APPLICANT: Schurigt, Hartmut Ernst
: APPLICANT: Tischert, Wilhelm
: APPLICANT: Wozny, Manfred
: TITLE OF INVENTION: Erythropoietin Conjugates
: FILE REFERENCE: Case 20805
: CURRENT APPLICATION NUMBER: US/10/014,363
: CURRENT FILING DATE: 2001-12-11
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3
: LENGTH: 174
: TYPE: PRT
: ORGANISM: CHO/dhfr-
: US-10-014-363-3

```

|                           |         |                  |        |                   |
|---------------------------|---------|------------------|--------|-------------------|
| Query Match               | 100.0%; | Score 846;       | DB 12; | Length 174;       |
| Best Local Similarity     | 100.0%; | Pred. No. 3e-84; |        |                   |
| Matches 165; Conservative | 0;      | Mismatches       | 0;     | Indels 0; Gaps 0; |

|    |     |  |     |
|----|-----|--|-----|
| Qy | 1   | APPRLODSRYLERYLLEAKAEENITTCGAHCGLNNITVPTDKVNFYAMKREMGQQA   | 60  |
| Db | 9   | APPRICDSRYLERYLLEAKAEENITTCGAHCGLNNITVPTDKVNFYAMKREMGQQA   | 68  |
| Qy | 61  | VEWMOGLALISEAVLRQALLVNSSQPWEPLQLHVDKAAVGLNSTITLLRALDAQKAIS | 120 |
| Db | 69  | VEWMOGLALISEAVLRQALLVNSSQPWEPLQLHVDKAAVGLNSTITLLRALDAQKAIS | 128 |
| Qy | 121 | PPDASAAPIRLTITADTFRKLLERYYSNPLGKTLKLYNGEACRRGD             | 165 |
| Db | 129 | PPDASAAPIRLTITADTFRKLLERYYSNPLGKTLKLYNGEACRRGD             | 173 |

```

US-10-014-363--5
: RESULT 7
: US-10-014-363--5
: Sequence 5, Application US/10014363
: Patent No. US2002011583A1
: GENERAL INFORMATION:
: APPLICANT: Burg, Josef
: APPLICANT: Engel, Alfired
: APPLICANT: Franze, Reinhard
: APPLICANT: Hilger, Bernd
: APPLICANT: Schürig, Hartmut Ernst
: APPLICANT: Tischer, Wilhelm
: APPLICANT: Wozny, Manfred
: TITLE OF INVENTION: Erythropoietin Conjugates
: FILE REFERENCE: Case 20805
: CURRENT APPLICATION NUMBER: US/10/014,363
: CURRENT FILING DATE: 2001-12-11
: NUMBER OF SEQ. ID NOS: 5
: SOFTWARE: patentIn version 3.1
: SEQ ID NO 5
: LENGTH: 174
: TYPE: PR1
: ORGANISM: CHO/dhfr-
: US-10-014-363--5

```

| Query Match | Similarity | 100.0%   | Score     | 846:       | DB | 12:    | Length | 174: |    |
|-------------|------------|--|-----------|------------|----|--------|--------|------|----|
| Best Local  | Similarity | 100.0%   | Pred. No. | 3e-84:     |    |        |        |      |    |
| Matches     | 165:       | Conservative   | 0:        | Mismatches | 0: | Indels | 0:     | Gaps | 0: |
| QY          | 1          | APPRIICDSRVLYERLLLEAKAEANITTTGCAEHCISLNIENITTVDTKYNFPAAMKRMVEYGOQA | 60        |            |    |        |        |      |    |
| Db          | 9          | APPRIICDSRVLYERLLLEAKAEANITTTGCAEHCISLNIENITTVDTKYNFPAAMKRMVEYGOQA | 68        |            |    |        |        |      |    |
| QY          | 61         | VEWVGGLAILSAVYLRGQALLVNSSQPMWPELDLHDKAVSGIRSLTTLLRALGAQKEAIS       | 120       |            |    |        |        |      |    |
| Db          | 69         | VEWVGGLAILSAVYLRGQALLVNSSQPMWPELDLHDKAVSGIRSLTTLLRALGAQKEAIS       | 128       |            |    |        |        |      |    |
| QY          | 121        | PPDAAASAAPLRTITADTFPKLLFRVYSNFKLTKLYLTGACPRGD                      | 165       |            |    |        |        |      |    |
| Db          | 129        | PPDAAASAAPLRTITADTFPKLLFRVYSNFKLTKLYLTGACPRGD                      | 173       |            |    |        |        |      |    |

```

/ RESULT 8
/ US-09-864-761-48502
/ Sequence 48502, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharron G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
/ FILE REFERENCE: Aecmca-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 48502
/ LENGTH: 60
/ TYPE: prt
/

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF053356.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
; OTHER INFORMATION: SWISSPROT HIT: P01588, EVALUO 9.00e-28
; OTHER INFORMATION: EST_HUMAN HIT: AA662379.1, EVALUO 3.00e-10
US-09-864-761-48502

Query Match
Best Local Similarity 100.0%; Score 294; DB 10; Length 60;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 VGGQAVVWGGALSLSEAVLRGQALLVNSQWPELQHVDAVSGLSLTLLALGAG 115
DB 1 VGGQAVVWGGALSLSEAVLRGQALLVNSQWPELQHVDAVSGLSLTLLALGAG 60

RESULT 9
US-09-975-063-2
; Sequence 2, Application US/09975063
; Patent No. US20020045255A1
; GENERAL INFORMATION:
; APPLICANT: POWELL, Jerry S.
; TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE, HIGH LEVEL
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dorsey & Whitney, LLP
; STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
; City: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.7/7.8
; CURRENT APPLICATION DATA: US/09/975.063
; FILING DATE: 10-OCT-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Roberts, Mark W.
; REGISTRATION NUMBER: 46,160
; REFERENCE/DOCKET NUMBER: 500582.03 (112893.109)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-903-8728
; TELEFAX: 206-903-8820
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-975-063-2

Query Match
Best Local Similarity 83.3%; Score 111; DB 10; Length 30;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 APRRLIDSRVLEKYLEAKAEENITGCA 30
DB 1 APRRLIDSRVLEKYLEAKAEENITGCA 30

RESULT 10
US-10-011-858-2
; Sequence 2, Application US/10011858
; Patent No. US20020137145A1
; GENERAL INFORMATION:
; APPLICANT: POWELL, Jerry S.
```

```

; TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE, HIGH LEVEL
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dorsey & Whitney, LLP
; STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
; City: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS-WORD
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/011,858
; FILING DATE: 05-NO. US20020137145A1-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Roberts, Mark W.
; REGISTRATION NUMBER: 46,160
; REFERENCE/DOCKET NUMBER: 500582.13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-903-8728
; TELEFAX: 206-903-8820
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-011-858-2

Query Match
Best Local Similarity 83.3%; Score 111; DB 12; Length 30;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 APRRLIDSRVLEKYLEAKAEENITGCA 30
DB 1 APRRLIDSRVLEKYLEAKAEENITGCA 30

RESULT 11
US-09-919-703-9
; Sequence 9, Application US/09919703
; Patent No. US20020165129A1
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; APPLICANT: Rabkin, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; FILE REFERENCE: 50216/003004
; CURRENT APPLICATION NUMBER: US/09/919,703
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 08/759,599
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-9

Query Match
Best Local Similarity 12.2%; Score 103; DB 9; Length 20;
```



Fri Jan 3 10:09:17 2003

us-09-853-731-1.rapb

Page 6

COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PC-DOS/MS-DOS  
CURRENT APPLICATION DATA: release #1.0, Version #1.25  
APPLICATION NUMBER: US/09/945,182  
FILING DATE: 31-Aug-2001  
CLASSIFICATION: <unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/808,324  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Lazar, Steven R.  
REGISTRATION NUMBER: 32,618  
REFERENCE/DOCKET NUMBER: 5202-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 498-8260  
TELEFAX: 617 876-5851  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 321 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-945-182-26

Query Match 8.5%; Score 72; DB 9; Length 321;  
Best Local Similarity 27.2%; Pred. No. 3.8; Mismatches 48; Indels 24; Gaps 3;  
Matches 31; Conservative 11;  
OY 57 GQQAWEVWQGLALSEAVLGRGALLVNSSQWPEPLQHVDRKAVSGLSRLTLLRALGAQK 116  
DB 77 GWEVFWVWQGL-----RHQPWKOLCTELRAAWGELDAGEAEARARAGPQ 120  
OY 117 EAISPPDAASAPLRTITADTFERKLFRVYSNFLRGKL-----KLYTGEACRTG 164  
DB 121 P--PPDLRLSLGFGRRVRPPQERALLVFTRSQKKNLFAEMREOLGSAAEAGPG 172

Search completed: January 2, 2003, 15:13:49  
Job time : 9.47432 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:10:05 ; Search time 13.9577 Seconds  
(without alignments)  
1136.446 Million cell updates/sec

Title: US-09-853-731-1

Perfect score: 846  
Sequence: 1 APRRLICDSVLEKRYLLEAK.....SNFLRGKLIKLTGEACTG 165

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR-73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 846   | 100.0       | 193    | 1 ZUHU   | erythropoietin pre |
| 2          | 764.5 | 90.4        | 192    | 1 J00173 | erythropoietin pre |
| 3          | 759.5 | 89.8        | 192    | 1 I84613 | erythropoietin pre |
| 4          | 713   | 84.3        | 188    | 1 I46083 | erythropoietin pre |
| 5          | 701   | 82.9        | 192    | 1 S28148 | erythropoietin pre |
| 6          | 685.5 | 81.0        | 194    | 1 I46401 | erythropoietin pre |
| 7          | 681   | 80.5        | 192    | 1 A24902 | erythropoietin pre |
| 8          | 680.5 | 80.4        | 195    | 1 UC7699 | erythropoietin pre |
| 9          | 678   | 80.1        | 190    | 2 I46199 | erythropoietin - r |
| 10         | 638   | 75.4        | 175    | 2 I46199 | erythropoietin - p |
| 11         | 602.7 | 90.0        | 353    | 2 G02729 | erythropoietin - d |
| 12         | 89    | 10.5        | 353    | 2 I80105 | thrombopoietin - h |
| 13         | 88    | 10.4        | 323    | 2 AB0323 | thrombopoietin pre |
| 14         | 87.5  | 10.3        | 346    | 2 AE0959 | ribonucleoside dip |
| 15         | 86    | 10.2        | 286    | 2 A55530 | Solute binding rec |
| 16         | 83    | 9.8         | 296    | 2 A10443 | megakaryocyte grow |
| 17         | 83    | 9.8         | 339    | 2 A83274 | probable 2-hydroxy |
| 18         | 80.5  | 9.5         | 3033   | 1 GNMVJ8 | UDP-N-acetylpyruvo |
| 19         | 79.5  | 9.4         | 1829   | 2 T35681 | genome polyprotein |
| 20         | 79    | 9.3         | 480    | 2 S56639 | probable sensory h |
| 21         | 78.5  | 9.3         | 813    | 2 AF0526 | ribosomal protein  |
| 22         | 78.5  | 9.3         | 897    | 2 A54656 | ATP-dependent heli |
| 23         | 78    | 9.2         | 448    | 2 T35450 | EGF receptor subst |
| 24         | 78    | 9.2         | 455    | 2 H97693 | ABC transporter AT |
| 25         | 78    | 9.2         | 455    | 2 AG2919 | methylamine utiliz |
| 26         | 77.5  | 9.2         | 747    | 1 S36741 | conserved hypotnet |
| 27         | 77.5  | 9.2         | 242    | 2 A51928 | probable copper-tr |
| 28         | 77    | 9.1         | 451    | 2 S75569 | hypothetical prote |
| 29         | 76.5  | 9.0         | 154    | 2 H82810 | bacterioferritin x |

|    |      |     |      |          |                    |
|----|------|-----|------|----------|--------------------|
| 30 | 76.5 | 9.0 | 425  | 2 AE3465 | mandelate racemase |
| 31 | 75.5 | 8.9 | 637  | 2 S75772 | hypothetical prote |
| 32 | 74.5 | 8.8 | 400  | 2 AB2922 | conserved hypotnet |
| 33 | 74.5 | 8.8 | 425  | 2 C97696 | res beta (AF305057 |
| 34 | 74.5 | 8.8 | 824  | 2 D64738 | ATP-dependent heli |
| 35 | 74   | 8.7 | 282  | 2 B37994 | RF2 protein - salm |
| 36 | 74   | 8.7 | 326  | 2 J04125 | thrombopoietin pre |
| 37 | 74   | 8.7 | 335  | 2 AH3625 | ribonucleoside dip |
| 38 | 74   | 8.7 | 1564 | 2 S55517 | probable transport |
| 39 | 73.5 | 8.7 | 401  | 2 H83911 | hypothetical prote |
| 40 | 73.5 | 8.7 | 476  | 1 S71789 | GCN5 protein - hum |
| 41 | 73.5 | 8.7 | 717  | 2 F82613 | VaG8 protein XFI98 |
| 42 | 73   | 8.6 | 263  | 2 B75361 | WD-repeat family p |
| 43 | 73   | 8.6 | 1089 | 2 S53978 | PSI1 protein - yea |
| 44 | 72.5 | 8.6 | 379  | 2 H69478 | NADH2 dehydrogenas |
| 45 | 72.5 | 8.6 | 401  | 2 AF3341 | precorrin-6y c5,15 |

#### ALIGNMENTS

##### RESULT 1

ZUHU  
erythropoietin precursor [validated] - human  
C:Species: Homo sapiens (man)  
C:Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 08-Dec-2000  
C:Accession: A01855; A24744; A25384; A22210; S56178  
R:Jacobs, K.; Shoemaker, C.; Ruderstorf, R.; Neill, S.D.; Kaufman, R.J.; Mufson, A.;  
Nature 313, 806-810, 1985  
A:Title: Isolation and characterization of genomic and cDNA clones of human erythro  
A:Reference number: A01855; MUID:85137899; PMID:3838366  
A:Accession: A01855  
A:Molecule type: mRNA; DNA  
A:Residues: 1-193 <JUC>  
A:Cross-references: GB:X02157; GB:X02158  
R:Lin, F.K.; Sugars, S.; Lin, C.H.; Browne, J.K.; Smalling, R.; Egrle, J.C.; Chen, K.K  
Proc. Natl. Acad. Sci. U.S.A. 82, 7580-7584, 1985  
A:Title: Cloning and expression of the human erythropoietin gene.  
A:Reference number: A24744; MUID:86067948; PMID:3865178  
A:Accession: A24744  
A:Molecule type: DNA  
A:Residues: 1-193 <LIN>  
A:Cross-references: GB:M11319; NID:g182197; PID:AA52400.1; PID:g182198  
R:Li, P.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser, E.  
J. Biol. Chem. 261, 3116-3121, 1986  
A:Title: Structural characterization of human erythropoietin.  
A:Reference number: A25384; MUID:86140080; PMID:3949763  
A:Accession: A25384  
A:Molecule type: protein  
A:Residues: 28-86, 'Q', 87-193 <LA1>  
A:Note: Experimental source: urine  
A:Note: forms without the carboxyl-terminal residue and the four carboxyl-terminal re  
R:Yanagawa, S.; Hirade, K.; Ohno, H.; Sasaki, R.; Chiba, H.; Ueda, M.; Goto, M.  
J. Biol. Chem. 259, 2707-2710, 1984  
A:Title: Isolation of human erythropoietin with monoclonal antibodies.  
A:Reference number: A22210; MUID:84135751; PMID:6698988  
A:Accession: A22210  
A:Molecule type: protein  
A:Residues: 28-29, 'X', 31-33, 'L', 35-50, 'X', 52-53, 'D', 55, 'G', 57 <YAN>  
R:Matsumoto, S.; Ikura, K.; Ueda, M.; Sasaki, R.  
Plant Mol. Biol. 27, 1163-1172, 1995  
A:Title: Characterization of a human glycoprotein (erythropoietin) produced in cultur  
A:Reference number: S56178; MUID:95284365; PMID:7766897  
A:Accession: S56178  
A:Molecule type: protein  
A:Residues: 28-33, 'X', 35-37 <MTS>  
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live  
C:Genetics:  
A:Gene: GDB:EPO  
A:Cross-references: GDB:119110; OMIM:133170  
A:Map position: 7q21.3-7q22.1  
A:introns: 5/1; 53/3; 82/3; 142/3  
C:Function:

A:Description: the primary inducer of erythrocyte formation  
C:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-193/Product: erythropoietin #status experimental <MAT>  
F:34-188,56-60/Disulfide bonds: #status experimental  
F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:153/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 846; DB 1; Length 193;  
Best Local Similarity 100.0%; Pred. No. 1,8e-74;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLLEAKENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 60  
DB 28 APPRLICDSRVLEERYLLLEAKENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 87  
OY 61 VEWMOGLALSEAVLRGQALLVNSQPMPERLQHDVKAAGSLRSLTTLRLALGAOKEAIS 120  
DB 88 VEWMOGLALSEAVLRGQALLVNSQPMPERLQHDVKAAGSLRSLTTLRLALGAOKEAIS 147  
OY 121 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRTGD 165  
DB 148 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRTGD 192

## RESULT 2

erythropoietin precursor - crab-eating macaque  
C:Species: Macaca fascicularis (crab-eating macaque)  
C:Date: 07-Sep-1990 #sequence\_revision 15-Nov-1996 #text\_change 22-Jun-1999  
C:Accession: J00173  
R:Lin, F.K.; Lin, C.H.; Lai, P.H.; Browne, J.K.; Egrile, J.C.; Smalling, R.; Fox, G.M.; C  
Gene 44, 201-209, 1986  
A:Title: Monkey erythropoietin gene: cloning, expression and comparison with the human  
A:Reference number: J00173; MUID:87055236; PMID:2877922  
A:Accession: J00173  
A:Molecule type: mRNA  
A:Residues: 1-192 <LIN>  
A:Cross-references: GB:M18189; GB:M15818; GB:M15819; GB:M18188; NID:9342093; PIDN:AAA368  
A:Experimental source: kidney  
C:Comment: This protein is the principal hormone involved in the regulation of erythrocy  
C:Function: Erythropoietin is produced by kidney or liver of adult mammals and by liver  
A:Description: the primary inducer of erythrocyte formation  
C:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-193/Product: erythropoietin #status experimental <MAT>  
F:34-188,56-60/Disulfide bonds: #status predicted  
F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:153/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 90.4%; Score 764.5; DB 1; Length 192;  
Best Local Similarity 91.5%; Pred. No. 1.4e-66;  
Matches 151; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLLEAKENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 60  
DB 28 APPRLICDSRVLEERYLLLEAKENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 87  
OY 61 VEWMOGLALSEAVLRGQALLVNSQPMPERLQHDVKAAGSLRSLTTLRLALGAOKEAIS 120  
DB 88 VEWMOGLALSEAVLRGQALLVNSQPMPERLQHDVKAAGSLRSLTTLRLALGAOKEAIS 146  
OY 121 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRTGD 165  
DB 147 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRTGD 191

RESULT 3  
184613  
erythropoietin precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 02-Aug-1996 #sequence\_revision 15-Nov-1996 #text\_change 22-Jun-1999  
C:Accession: I84613

R:Wen, D.; Boissel, J.  
Blood 82, 1507-1516, 1993

A:Title: Erythropoietin structure-function relationships: High degree of sequence hom  
A:Reference number: I46083; MUID:93372347; PMID:8364201

A:Accession: I84613  
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1-192 <RES>

A:Cross-references: GB:I10609; NID:9342095; PIDN:AAA36842.1; PID:9342096

C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live

C:Function: the primary inducer of erythrocyte formation

A:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver

F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-192/Product: erythropoietin #status predicted <MAT>  
F:34-187,56-60/Disulfide bonds: #status predicted

F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:153/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 89.8%; Score 759.5; DB 1; Length 192;  
Best Local Similarity 90.3%; Pred. No. 4.3e-66;  
Matches 149; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLLEAKENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 60  
DB 28 APPRLICDSRVLEERYLLLEAKENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 87  
OY 61 VEWMOGLALSEAVLRGQALLVNSQPMPERLQHDVKAAGSLRSLTTLRLALGAOKEAIS 120  
DB 88 VEWMOGLALSEAVLRGQALLVNSQPMPERLQHDVKAAGSLRSLTTLRLALGAOKEAIS 146  
OY 121 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRTGD 165  
DB 147 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRTGD 191

## RESULT 4

erythropoietin precursor - cat (fragment)  
C:Species: Felis silvestris catus (domestic cat)  
C:Date: 16-Aug-1996 #sequence\_revision 15-Nov-1996 #text\_change 22-Jun-1999  
C:Accession: I46083  
R:Wen, D.; Boissel, J.  
Blood 82, 1507-1516, 1993  
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom  
A:Reference number: I46083; MUID:93372347; PMID:8364201  
A:Accession: I46083  
A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1-188 <WEN>

A:Cross-references: GB:I10606; NID:9163820; PIDN:AAA30807.1; PID:9163821

C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live

C:Function: the primary inducer of erythrocyte formation

A:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver

F:1-22/Domain: signal sequence (fragment) #status predicted <SIG>  
F:23-188/Product: erythropoietin #status predicted <MAT>  
F:29-183,51-55/Disulfide bonds: #status predicted

F:46,60,105/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:148/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 84.3%; Score 713; DB 1; Length 188;  
Best Local Similarity 84.2%; Pred. No. 1.3e-61;  
Matches 139; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLLEAKENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 60  
DB 23 APPRLICDSRVLEERYLLLEAKENITTCAGHCISLNETITVPDTKVNFAKRMVEVGOQA 82





|    |     |  |     |
|----|-----|--|-----|
| Db | 24  | APR--ACDLRVLSTLRRSHVLSKLSGCEVHPLPPVLLPAVDLSGEKTKTQMEETKA   | 81  |
| OY | 61  | VEWMOGLALISEAVL--RGQALLVNSSOPWEPDLDHDKAVSGLSRLTTLRALGAQKEA | 118 |
| Db | 82  | QDILGAVTLLEGWARGQGLGPTCLSLSLQSLSEOVRLILGALQSL-----LGTQ---  | 132 |
| OY | 119 | ISPPDASAAPLRTITADTFEKLFRVYSNPLRGKL                         | 154 |
| Db | 133 | -LPDQ-----RTTAHKDPNATFLSFQHLNKGVR                          | 161 |

RESULT 12  
180105  
thrombopoietin precursor - human  
N:Alternate names: c-MPL ligand; megakaryocyte growth and development factor precursor  
C:Species: Homo sapiens (man)  
C:Date: 24-May-1996 #sequence, revision 24-May-1996 #text\_change 20-Jun-2000  
C:Accession: I59281; 180105; S45331; S48740; 138672; 152610  
R:Poster, D.C.; Sprecher, C.A.; Grant, F.J.; Kramer, J.M.; Kilper, J.L.; Holly, R.D.; W  
Proc. Natl. Acad. Sci. U.S.A. 91, 13023-13027, 1994  
A:Title: Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosomal  
A:Reference number: I59281; MUID:95108091; PMID:7809166  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-353 <RE2>  
A:Cross-references: GB:I36051; NID:g533214; PIDN:AAC37568.1; PID:g533215  
A:Accession: 180105  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-353 <RES>  
A:Cross-references: GB:I36052; NID:g533216; PIDN:AAC37566.1; PID:g533217  
R:de Sauvage, F.J.; Hass, P.E.; Spencer, S.D.; Malloy, B.E.; Gurney, A.L.; Spencer, S.A.  
D.V.; Eaton, D.L.  
Nature 369, 533-538, 1994  
A:Title: Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl ligand.  
A:Reference number: S45331; MUID:94261202; PMID:8202154  
A:Accession: S45331  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-353 <SAU>  
A:Cross-references: GB:I33410; NID:g506826; PIDN:AAA59857.1; PID:g506827  
R:Somma, Y.; Akahori, H.; Seki, N.; Horl, T.; Ogami, K.; Kato, T.; Shlnada, Y.; Kawamura  
FEBS Lett. 353, 57-61, 1994  
A:Title: Molecular cloning and chromosomal localization of the human thrombopoietin gene  
A:Reference number: S48740; MUID:95010765; PMID:7926023  
A:Accession: S48740  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-353 <SOH>  
A:Cross-references: GB:I32046; NID:g577319; PIDN:BAA06807.1; PID:g577320  
R:Barclay, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.S.;  
Cell 77, 1117-1124, 1994  
A:Title: Identification and cloning of a megakaryocyte growth and development factor tha  
A:Reference number: A54463; MUID:94291201; PMID:8020099  
A:Accession: 138672  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-112, 'E', 114-353 <RE3>  
A:Cross-references: EMBL:U011025; NID:g511223; PIDN:AAA50553.1; PID:g558078  
R:Gurney, A.L.; Kiang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.  
Blood 85, 981-988, 1995  
A:Title: Genomic structure, chromosomal localization, and conserved alternative splice f  
A:Reference number: I52610; MUID:95152076; PMID:7849319  
A:Accession: 152610  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-353 <RE4>  
A:Cross-references: GB:S76771; NID:g914225; PIDN:AAB33390.1; PID:g914226  
C:Genetics:  
A:Gene: GDB:THPO, MGDP  
A:Cross-references: GDB:374007; OMIM:600044

```

A:Map position: 3q26.3-3q427
A:introns: 5/1: 47/3: 76/3: 132/3
C:keywords: alternative splicing; cytokine; glycoprotein

Query Match      10.5%      Score 89; DB 2; Length 353;
Best Local Similarity 26.3%; Pred. NO. 0.75;
Matches 41; Conservative 20; Mismatches 75; Indels 20; Gaps 5;

```

[illegible]

RESULT 13  
AB0323  
ribonucleoside-diphosphate reductase (EC 1.17.4.1) beta chain [imported] - Yersinia p  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AB0323  
J:Parkhill, J., Wren, B.W.; Thomson, N.R.; Taitball, R.W.; Holden, M.T.G.; Prentice, M.  
deno-Tarragge, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.  
L., M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AB0323  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-323 <KUP>  
A:Cross-references: GB:AL590842; PIDN:CAC92889.1; PID:G15980631; GSPDB:GN00175  
C:Genetics:  
A:Gene: nrdf  
C:Superfamily: ribonucleoside-diphosphate reductase beta  
C:Keywords: oxidoreductase

[illegible]

A:Reference number: AB0502; PMID:11677608  
A:Accession: AE0959  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1346 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD03169.1; PID:g16504804; GSPDB:GN00176  
C:Genetics: STY3952

Query Match 10.3%; Score 87.5; DB 2; Length 346;  
Best Local Similarity 26.7%; Pred. No. 1;  
Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

OY 10 RYLERYLEKAEKNITTG--CAEHCISNE--NIIVPDTRKVNFTYANKRMEVGOAVEVWQ 65  
Db 217 RNLDQELERHPDANVAGSAIAEAAGRGRLTTLIVSTL-----THQVTR 267  
OY 66 GLALLSEAVLRGOALLVNSO--PMEPLQHLVDAVSGLSLTLLRALGAQ--KEAISP 122  
Db 268 GLK-----RGHILALSDQAMQ-----GELAITQSTKYLOGQVPEINISPP 309  
OY 123 -----DASAAPLRTITADPFRKLPVYSNFLRGKALKYGEA 160  
Db 310 VILITNNNADSAVRRSLSPPGFRPY-----LYQITSEA 344

RESULT 15  
A55530  
megakaryocyte growth and development factor, long form - human  
N:Alternate names: MPL ligand, long form  
C:Species: Homo sapiens (man)  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
C:Accession: A55530  
R:Chang, M.; McIninch, J.; Basu, R.; Shuttler, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs, S.  
J. Biol. Chem. 270, 511-514, 1995  
A:Title: Cloning and characterization of the human megakaryocyte growth and development  
A:Reference number: A55530; MUID:95122483; PMID:7822271  
A:Accession: A55530  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-286 <CHA>  
A:Cross-references: GB:U17071  
C:Genetics: MGDF  
A:Map position: 3q26.3  
C:Keywords: alternative splicing; cytokine

Query Match 10.2%; Score 86; DB 2; Length 286;  
Best Local Similarity 26.6%; Pred. No. 1.1;  
Matches 41; Conservative 18; Mismatches 75; Indels 20; Gaps 5;  
OY 1 APPRLICDSRVLEKYLEKAEKNITTGCAEHCISNENITVDPDTRKVNFTYANKRMEVGOA 60  
Db 24 APP--ACDLRVLSKLLRDSHVLSRSLSCPEVHPLPTVLPAVDFSGEWTQMEETKA 81  
OY 61 VEVWQGLALLSEAVL--RGOALLVNSQSPWEPLQHLVDAVSGLSLTLLRALGAQKEA 118  
Db 82 ODILGAVTLLEGVMAARGQLGPTCTLSLGLQSGQVRLLLGALQSL-----LQTV--- 132  
OY 119 ISPPDASAAPLRTITADPFRKLPVYSNFLRGK 152  
Db 133 -LPFG-----RTVAHKDPNAIFLSFQHLLRGK 159

Search completed: January 2, 2003, 15:13:26  
Job time : 14.9577 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:07:09 : Search time 8 47432 Seconds  
(without alignments)  
807.569 Million cell updates/sec

Title: US-09-853-731-1

Perfect score: 846  
Sequence: 1 APPRLICDSRVLEERYLEAK.....SNFLRGKIKITGECACRTGD 165

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 846   | 100.0       | 193    | 1  | EPO_HUMAN   |
| 2          | 764.5 | 90.4        | 192    | 1  | EPO_MACFA   |
| 3          | 759.5 | 89.8        | 192    | 1  | EPO_MACMU   |
| 4          | 706   | 83.5        | 192    | 1  | EPO_FETCA   |
| 5          | 701   | 82.9        | 192    | 1  | EPO_RAT     |
| 6          | 692.5 | 81.9        | 192    | 1  | EPO_BOVIN   |
| 7          | 689   | 81.4        | 192    | 1  | EPO_MOUSE   |
| 8          | 685.5 | 81.0        | 194    | 1  | EPO_SHEEP   |
| 9          | 678   | 80.1        | 190    | 1  | EPO_PIG     |
| 10         | 638   | 75.4        | 175    | 1  | EPO_CANFA   |
| 11         | 109   | 12.9        | 352    | 1  | TPO_CANFA   |
| 12         | 89    | 10.5        | 353    | 1  | TPO_HUMAN   |
| 13         | 80.5  | 9.5         | 3033   | 1  | POLG_HCVJ8  |
| 14         | 78.5  | 9.3         | 897    | 1  | EPI5_MOUSE  |
| 15         | 78    | 9.2         | 747    | 1  | ATCS_SYNPF  |
| 16         | 77    | 9.1         | 548    | 1  | CH60_BUCAP  |
| 17         | 76.5  | 9.0         | 386    | 1  | CTBP_DROME  |
| 18         | 76    | 9.0         | 551    | 1  | CH60_BUCAP  |
| 19         | 74.5  | 8.8         | 552    | 1  | CH60_PEST   |
| 20         | 74.5  | 8.8         | 809    | 1  | HRPB_ECOLI  |
| 21         | 74    | 8.7         | 326    | 1  | TPO_RAT     |
| 22         | 74    | 8.7         | 1564   | 1  | PDR1_YEAST  |
| 23         | 73.5  | 8.7         | 830    | 1  | GCI2_MOUSE  |
| 24         | 73.5  | 8.7         | 837    | 1  | GCI2_HUMAN  |
| 25         | 73    | 8.6         | 263    | 1  | YH25_DEIRA  |
| 26         | 73    | 8.6         | 1089   | 1  | YH25_YEAST  |
| 27         | 72.5  | 8.6         | 762    | 1  | SLAP_ACEKI  |
| 28         | 72    | 8.5         | 356    | 1  | TPO_MOUSE   |
| 29         | 72    | 8.5         | 548    | 1  | CH60_BUCAI  |
| 30         | 70.5  | 8.3         | 217    | 1  | YHND_ECOLI  |
| 31         | 70.5  | 8.3         | 381    | 1  | MODD_MYCAV  |
| 32         | 70.5  | 8.3         | 543    | 1  | CH60_BARBA  |
| 33         | 70.5  | 8.3         | 896    | 1  | EPI5_HUMAN  |

## ALIGNMENTS

| RESULT 1 | ID   | EPO_HUMAN                         | STANDARD: | PRT: | 193 AA. |
|----------|--|-----------------------------------|-----------|------|---------|
| AC       | P01588   | Q9UHA0; Q9UEZ5; Q9UD20;           |           |      |         |
| DT       | 21-JUL-1986  | (Rel. 01, Created)                |           |      |         |
| DT       | 21-JUL-1986  | (Rel. 01, last sequence update)   |           |      |         |
| DE       | 16-OCT-2001  | (Rel. 40, last annotation update) |           |      |         |
| DE       | Erythropoietin precursor (Epoetin).                                    |                                   |           |      |         |
| GN       | EPO.   |                                   |           |      |         |
| OS       | Homo sapiens (Human).  |                                   |           |      |         |
| OC       | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |                                   |           |      |         |
| OC       | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.              |                                   |           |      |         |
| OX       | NCBI_Taxid:9606;   |                                   |           |      |         |
| RN       | [1]  |                                   |           |      |         |
| RP       | SEQUENCE FROM N.A.   |                                   |           |      |         |
| RX       | MEDLINE-85137899; PubMed-3838366;                                      |                                   |           |      |         |
| RA       | Jacobs K., Shoemaker C., Ruderstorf R., Neill S.D., Kaufman R.J.,      |                                   |           |      |         |
| RA       | Mufson A., Seehra J., Jones S.S., Hewick R., Fritsch E.F.,             |                                   |           |      |         |
| RA       | Kawakita M., Shimizu T., Miyake T.;                                    |                                   |           |      |         |
| RT       | Isolation and characterization of genomic and cDNA clones of human     |                                   |           |      |         |
| RT       | erythropoietin."   |                                   |           |      |         |
| RL       | Nature 313:806-810(1985).  |                                   |           |      |         |
| RN       | [2]  |                                   |           |      |         |
| RP       | SEQUENCE FROM N.A.   |                                   |           |      |         |
| RX       | MEDLINE-86067948; PubMed-3865178;                                      |                                   |           |      |         |
| RA       | Lin F.-K., Suggs S., Lin C.-H., Browne J.K., Egitte J.C.,              |                                   |           |      |         |
| RA       | Chen K.K., Fox G.M., Martin F., Stabinsky Z., Bedrawi S.M., Lai P.-H., |                                   |           |      |         |
| RA       | Goldwasser E.;   |                                   |           |      |         |
| RT       | "Cloning and expression of the human erythropoietin gene.";            |                                   |           |      |         |
| RT       | Proc. Natl. Acad. Sci. U.S.A. 82:7580-7584(1985).                      |                                   |           |      |         |
| RN       | [3]  |                                   |           |      |         |
| RP       | SEQUENCE FROM N.A.   |                                   |           |      |         |
| RX       | MEDLINE-99018118; PubMed-9799793;                                      |                                   |           |      |         |
| RA       | Gloekner G., Scherer S., Schattevoy R., Boright A., Weber J.,          |                                   |           |      |         |
| RA       | Tsui L.-C., Rosenthal A.;  |                                   |           |      |         |
| RT       | "Large-scale sequencing of two regions in human chromosome 7q22:       |                                   |           |      |         |
| RT       | analysis of 650 kb of genomic sequence around the EPO and CDT1 loci    |                                   |           |      |         |
| RT       | reveals 17 genes.";  |                                   |           |      |         |
| RL       | Genome Res. 8:1060-1073(1998).   |                                   |           |      |         |
| RN       | [4]  |                                   |           |      |         |
| RP       | SEQUENCE FROM N.A.   |                                   |           |      |         |
| RA       | Rupert J.L., Hochachka P.W.;   |                                   |           |      |         |
| RT       | "Erythropoietin gene sequence in the Quechua, a high altitude native   |                                   |           |      |         |
| RT       | population.";  |                                   |           |      |         |
| RL       | Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.                |                                   |           |      |         |
| RN       | [5]  |                                   |           |      |         |
| RP       | SEQUENCE OF 58-193 FROM N.A., AND VARIANTS HEPATOCYLLULAR CARCINOMA.   |                                   |           |      |         |
| RX       | MEDLINE-93384593; PubMed-8396923;                                      |                                   |           |      |         |
| RA       | Funakoshi A., Muta H., Baba T., Shimizu S.;                            |                                   |           |      |         |
| RT       | "Gene expression of mutant erythropoietin in hepatocellular            |                                   |           |      |         |
| RT       | carcinoma.";   |                                   |           |      |         |
| RL       | Biochem. Biophys. Res. Commun. 195:717-722(1993).                      |                                   |           |      |         |
| RN       | [6]  |                                   |           |      |         |
| RP       | SEQUENCE OF 28-193, AND DISULFIDE BONDS.                               |                                   |           |      |         |
| RT       | TISSUE-Urine;  |                                   |           |      |         |
| RC       | MEDLINE-86140080; PubMed-3949763;                                      |                                   |           |      |         |

|    |      |     |      |   |            |
|----|------|-----|------|---|------------|
| 34 | 70   | 8.3 | 319  | 1 | RIR4_SALTY |
| 35 | 70   | 8.3 | 547  | 1 | CH60_LEGPN |
| 36 | 69.5 | 8.2 | 907  | 1 | GACS_PSEST |
| 37 | 69.5 | 8.2 | 1327 | 1 | TNR1_HUMAN |
| 38 | 69   | 8.2 | 353  | 1 | NADA_YERPE |
| 39 | 69   | 8.2 | 544  | 1 | CH60_FRATU |
| 40 | 69   | 8.2 | 552  | 1 | CH60_COXBU |
| 41 | 69   | 8.2 | 1128 | 1 | DNR1_HSVSA |
| 42 | 68.5 | 8.1 | 360  | 1 | CYS2_MAIZE |
| 43 | 68.5 | 8.1 | 418  | 1 | GUN2_TTRIE |
| 44 | 68.5 | 8.1 | 778  | 1 | RG12_MOUSE |
| 45 | 68   | 8.0 | 366  | 1 | QVEA_SYNY3 |

|        |              |
|--------|--------------|
| P17424 | salmonella   |
| P26878 | leptomonella |
| P48027 | pseudomonas  |
| O95271 | homo sapien  |
| O82978 | yersinia pe  |
| P94798 | franciseilla |
| P19421 | coxiella bu  |
| P24910 | herpesvirus  |
| O10717 | zea mays (m  |
| P07982 | trichoderma  |
| O61193 | mus musculus |
| O58850 | synecocyst   |

RA Lai P.H.,Everett R., Wang F.F., Arkawa T., Goldasser E.,  
RT "Structural characterization of human erythropoietin.";  
RN J. Biol. Chem. 261:3116-3121(1986).  
RX PRELIMINARY SEQUENCE OF 28-57  
RX MEDLINE=84135751; Pubmed=696989;  
RA Yanagawa S., Hirade K., Ohmota H., Sasaki R., Chiba H., Ueda M.,  
RT "Isolation of human erythropoietin with monoclonal antibodies.";  
RN J. Biol. Chem. 259:2707-2710(1984).  
RP STRUCTURE OF CARBOHYDRATES.  
RX MEDLINE=88153657; Pubmed=3346214;  
RA Takeuchi M., Takasaki S., Miyazaki H., Kato T., Hoshi S., Kochibe N.,  
RT "Comparative study of the asparagine-linked sugar chains of human  
RT erythropoietins purified from urine and the culture medium of  
RT recombinant Chinese hamster ovary cells.";  
RN J. Biol. Chem. 263:3657-3663(1988).  
RP STRUCTURE OF CARBOHYDRATES.  
RX MEDLINE=89118279; Pubmed=3219367;  
RA Sasaki H., Ochi N., Dell A., Fukuda M.;  
RT "Site-specific glycosylation of human recombinant erythropoietin:  
RT analysis of glycoproteins or peptides at each glycosylation site by  
RT fast-atom bombardment mass spectrometry.";  
RN J. Biochemistry 27:8618-8626(1988).  
RP STRUCTURE OF CARBOHYDRATES.  
RX MEDLINE=92314463; Pubmed=1820196;  
RA Takeuchi M., Kobata A.;  
RT "Structures and functional roles of the sugar chains of human  
RT erythropoietins.";  
RN Glycobiology 1:337-346(1991).  
X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX MEDLINE=98445092; Pubmed=9774108;  
RA Zhan R.S., Reid S.W., Li C., Cheetham J.C., Aoki K.H., Liu B.,  
RA Sheng H., Osslund T.D., Chirino A.D., Zhang J., Finer-Moore J.,  
RA Elliott S., Stiney K., Katz B.A., Matthews D.J., Wendoloski J.J.,  
RT "Efficiency of signalling through cytokine receptors depends  
RT critically on receptor orientation.";  
RN Nature 395:511-516(1998).  
RL -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE  
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A  
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.  
CC -1- SUBCELLULAR LOCATION: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS  
CC -1- TISSUE SPECIFICITY: SECRETED.  
CC -1- AND BY LIVER OF FETAL OR NEONATAL MAMMALS.  
CC PHARMACEUTICAL: Used for the treatment of anemia. Available under  
CC the names Epogen (Amgen), Epogin (Chugai), Epomax (Eliane), Eprex  
CC (Janssen-cilag), Neorecomon or Recormon (Roche), and Procrit  
CC (Ortho Biotech). Variations in the glycosylation pattern of EPO  
CC distinguishes these products. Epogen, Epogin, Eprex and Procrit  
CC are genetically known as epoetin alfa, Neorecomon and Recormon as  
CC epoetin beta and Epomax as epoetin omega.  
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.  
CC -1- DATABASE: NAME=Red Systems/ cytokine source book. EPO:  
CC WWW="http://www.rndsystems.com/asp/g.steblidder.aspbodyId=197".  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch))  
CC -----  
DR EMBL; X02158; CAA26095.1; -;  
DR EMBL; X02157; CAA26094.1; -;  
DR EMBL; M1319; AAA52400.1; -;  
DR EMBL; AF053356; AAC878791.1; -;

| ID                    | EPO_MACFA  | STANDARD; | PRT;               | 192 AA.              |
|-----------------------|--|-----------|--------------------|----------------------|
| DR                    | EMBL: AF202308; AAF23132.1; -  | JOINED.   |                    |                      |
| DR                    | EMBL: AF202306; AAF23132.1; -  | JOINED.   |                    |                      |
| DR                    | EMBL: AF202307; AAF23132.1; -  | JOINED.   |                    |                      |
| DR                    | EMBL: AF202310; AAF23132.1; -  | JOINED.   |                    |                      |
| DR                    | EMBL: AF202309; AAF23132.1; -  | JOINED.   |                    |                      |
| DR                    | EMBL: AF202311; AAF23132.1; -  | JOINED.   |                    |                      |
| DR                    | EMBL: AF202311; AAF23132.1; -  | JOINED.   |                    |                      |
| DR                    | EMBL: AF202311; AAF23132.1; -  | JOINED.   |                    |                      |
| DR                    | EMBL: AF202313; AAF23134.1; -  | JOINED.   |                    |                      |
| DR                    | EMBL: S64548; AAF23134.1; -  | JOINED.   |                    |                      |
| DR                    | PIR: A01955; Z0H0.   |           |                    |                      |
| DR                    | PIR: A21584; A23584.   |           |                    |                      |
| DR                    | PIR: A24744; A25744.   |           |                    |                      |
| DR                    | PDB: A22210; A22210.   |           |                    |                      |
| DR                    | PDB: 1EEK; 01-OCT-99.  |           |                    |                      |
| DR                    | Glycosulteds; P01588; -  |           |                    |                      |
| DR                    | Gnew; HGNC:3415; EPO.  |           |                    |                      |
| DR                    | MIM; 133170; -   |           |                    |                      |
| DR                    | InterPro; IPR001323; EPO_TPO.  |           |                    |                      |
| DR                    | InterPro; IPR003013; Erythroptn.   |           |                    |                      |
| DR                    | PRINTS; P00758; EPO_TPO; 1.  |           |                    |                      |
| DR                    | PRINTS; P00272; ERYTHROPTN.  |           |                    |                      |
| DR                    | PROSITE; PS00817; EPO_TPO; 1.  |           |                    |                      |
| KM                    | Erythrocyte maturation; Glycoprotein; Hormone; Signal; Pharmaceutical; 3D structure. |           |                    |                      |
| KM                    | Signal.  |           |                    |                      |
| FT                    | CHAIN  | 1         | 27                 |                      |
| FT                    | PROPE  | 28        | 193                |                      |
| FT                    | PROPE  | 190       | 193                |                      |
| FT                    | DISULFID   | 34        | 188                |                      |
| FT                    | DISULFID   | 36        | 60                 |                      |
| FT                    | CARBOHYD   | 51        | 51                 |                      |
| FT                    | CARBOHYD   | 65        | 65                 |                      |
| FT                    | CARBOHYD   | 110       | 110                |                      |
| FT                    | CARBOHYD   | 153       | 153                |                      |
| FT                    | VARIANT  | 131       | 132                |                      |
| FT                    | VARIANT  | 149       | 149                |                      |
| FT                    | CONFLICT   | 40        | 40                 |                      |
| FT                    | CONFLICT   | 85        | 85                 |                      |
| FT                    | CONFLICT   | 140       | 140                |                      |
| FT                    | SEQUENCE   | 193 AA;   | 21306 MW;          |                      |
| Query Match           |  | 100.0%;   | Score 846;         | DB 1; Length 193;    |
| Best Local Similarity |  | 100.0%;   | Pred. No. 1.9e-76; |                      |
| Matches 165;          | Conservative   | 0;        | Mismatches         | 0; Indels 0; Gaps 0; |
| QY                    | 1 APPRICDSRVRYLLLEAKAEINNTTGCAGHCISLNIETVVDPTKYNFAAMRMEVGOA 60                       |           |                    |                      |
| DB                    | 28 APPRICDSRVRYLLLEAKAEINNTTGCAGHCISLNIETVVDPTKYNFAAMRMEVGOA 87                      |           |                    |                      |
| QY                    | 61 VEVWOGALLLSFAVIRGOALLVNSSQWPEDLQHDRAVSGRSITLTLRALGAQKEALS 120                     |           |                    |                      |
| DB                    | 88 VEVWOGALLLSFAVIRGOALLVNSSQWPEDLQHDRAVSGRSITLTLRALGAQKEALS 147                     |           |                    |                      |
| QY                    | 121 PPDASAAPLRTITADTFKRLFRVSNFRLGKLTLYTGACRTGD 165                                   |           |                    |                      |
| DB                    | 148 PPDASAAPLRTITADTFKRLFRVSNFRLGKLTLYTGACRTGD 192                                   |           |                    |                      |
| RESULT 2              |  |           |                    |                      |
| EPO_MACFA             | STANDARD;  | PRT;      | 192 AA.            |                      |
| AC                    | P07865;  |           |                    |                      |
| DT                    | 01-AUG-1988 (Rel. 08, Created)   |           |                    |                      |
| DT                    | 01-AUG-1988 (Rel. 08, Last sequence update)  |           |                    |                      |
| DT                    | 16-OCT-2001 (Rel. 40, Last annotation update)  |           |                    |                      |
| DE                    | Erythropoietin precursor.  |           |                    |                      |



```

GN Epo.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87053236; PubMed=2877922;
RA Lin E.-K., Lin C.-H., Lai P.-H., Browne J.K., Egrie J.C., Smalling R.,
RA Fox G.M., Chen K.K., Castro M., Suggs S.;
RT "Monkey erythropoietin gene: cloning, expression and comparison with
RT the human erythropoietin gene."
RL Gene 44:201-209(1986).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE Epo / TPO FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M18189; AAA36841.1; -.
DR PIR; J00173; J00173.
DR HSSP; P01588; ICN4.
DR InterPro; IPR001323; Epo_TPO.
DR InterPro; IPR003013; Erythropo.
DR Pfam; PF00758; Epo_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; Epo_TPO; 1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 27
FT CHAIN 28 192
FT DISULFID 34 187
FT CARBOHYD 51 51
FT CARBOHYD 65 65
FT CARBOHYD 110 110
FT CARBOHYD 152 152
SQ SEQUENCE 192 AA; 21113 MW; E8A900F442AD4522 CRC64;

Query Match 90.4%; Score 764.5; DB 1; Length 192;
Best Local Similarity 91.5%; Pred. No. 2,2e-68;
Matches 151; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 1 APRRLICDSRYVERLYLLEKAEKNITTCGAECISLNNITVPTKYNFAAKRMKGGOA 60
DB 28 APRRLICDSRYVERLYLLEKAEKNITTCGAECISLNNITVPTKYNFAAKRMKGGOA 87
QY 61 VEVWQGLALISEAVLRGQAVLNVSSQPEPQLQHVDAVSGLSLTTLLRALGAQKEALS 120
DB 88 VEVWQGLALISEAVLRGQAVLNVSSQPEPQLQHVDAVSGLSLTTLLRALGAQ- EALS 146
QY 121 PPDASAAPLRITTTADTFKRLFRVYSNPLRGKILKLTGTGACRTGD 165
DB 147 LPDASAAPLRITTTADTFKRLFRVYSNPLRGKILKLTGTGACRRGD 191

RESULT 3
EPO_MACMU
ID EPO_MACMU STANDARD: PRT; 192 AA.
AC Q28513;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Erythropoietin precursor.
GN Epo.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93372347; PubMed=8364201;
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals."
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE Epo / TPO FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L10609; AAA36842.1; -.
DR HSSP; P01588; ICN4.
DR InterPro; IPR001323; Epo_TPO.
DR InterPro; IPR003013; Erythropo.
DR Pfam; PF00758; Epo_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; Epo_TPO; 1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 27
FT CHAIN 28 192
FT DISULFID 34 187
FT CARBOHYD 51 51
FT CARBOHYD 65 65
FT CARBOHYD 110 110
FT CARBOHYD 152 152
SQ SEQUENCE 192 AA; 21081 MW; 275560A264628CD1 CRC64;

Query Match 89.8%; Score 759.5; DB 1; Length 192;
Best Local Similarity 90.3%; Pred. No. 6.7e-68;
Matches 149; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

QY 1 APRRLICDSRYVERLYLLEKAEKNITTCGAECISLNNITVPTKYNFAAKRMKGGOA 60
DB 28 APRRLICDSRYVERLYLLEKAEKNITTCGAECISLNNITVPTKYNFAAKRMKGGOA 87
QY 61 VEVWQGLALISEAVLRGQAVLNVSSQPEPQLQHVDAVSGLSLTTLLRALGAQKEALS 120
DB 88 VEVWQGLALISEAVLRGQAVLNVSSQPEPQLQHVDAVSGLSLTTLLRALGAQ- EALS 146
QY 121 PPDASAAPLRITTTADTFKRLFRVYSNPLRGKILKLTGTGACRTGD 165
DB 147 LPDASAAPLRITTTADTFKRLFRVYSNPLRGKILKLTGTGACRRGD 191

RESULT 4
EPO_FELCA
ID EPO_FELCA STANDARD: PRT; 192 AA.
AC P33708;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)

```

```

DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Goodman R.E., Bell R.G.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE OF 5-192 FROM N.A.
RX MEDLINE=93372347; PubMed=8364201;
RA Men D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
sequence homology among mammals."
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
-----
DR EMBL; U00685; AAA18282.1; -
DR EMBL; L10606; AAA30807.1; -
DR HSSP; P01588; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropo.
DR Pfam; PF00758; EPO_TPO.1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO.1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 26
FT CHAIN 27 192
FT DISULFID 33 187
FT DISULFID 55 187
FT CARBOHYD 50 59
FT CARBOHYD 64 59
FT CARBOHYD 109 109
FT CONFILCT 44 44
SQ SEQUENCE 192 AA; 20914 MW; 61C5EA0FE5E937293 CRC64;

Query Match 83.5%; Score 706; DB 1; length 192;
Best Local Similarity 83.6%; Pred. No. 1,36-62;
Matches 138; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

```

```

EPO_RAT
ID EPO_RAT STANDARD; PRT; 192 AA.
AC P29676; P70504;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=Mistral; TISSUE=Kidney;
RX MEDLINE=93042015; PubMed=1420369;
RA Nagao M., Suga H., Okano M., Masuda S., Narita H., Ikura K.,
RA Sasaki R.;
RT Nucleotide sequence of rat erythropoietin."
RL Biochim. Biophys. Acta 1171:99-102(1992).
RN 12
RP SEQUENCE OF 4-192 FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=93372347; PubMed=8364201;
RA Men D., Boissel J.P.R., Tracy T.E., Mulcahy L.S., Czelusniak J.,
RA Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
sequence homology among mammals."
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
-----
DR EMBL; D10763; BAA01593.1; -
DR EMBL; L10608; BAA41126.1; -
DR PIR; S28148; S28148.
DR HSSP; P01588; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropo.
DR Pfam; PF00758; EPO_TPO.1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO.1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 26
FT CHAIN 27 192
FT DISULFID 33 187
FT DISULFID 50 59
FT CARBOHYD 64 64
FT CARBOHYD 109 109
SQ SEQUENCE 192 AA; 21286 MW; 3EA632737E7D2443 CRC64;

Query Match 82.9%; Score 701; DB 1; length 192;
Best Local Similarity 82.4%; Pred. No. 4e-62;
Matches 136; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

```

```

Db 87 VEWVQGSJLSLSEALLQALQANSSQSPESJQLHDKAIISGJRSJFSLRLVJGAQKELMS 146
Oy 121 PPDAASAPLRTITADPFRKLFRRYSNFKRKLKLYGCAERTGD 165
Db 147 PPDAATQAAPLRTITADPFCCKLFRRYSNFKRKLKLYGCAERGD 191

RESULT 6
ID_EPO_BOVIN STANDARD; PRT; 192 AA.
AC P48617;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A.
RP STRAIN=Bozan. TISSUE=Kidney;
RC MEDLINE=96257233; PubMed=8666286;
RA Sulman H.B., Majlwa P.A.O., Feldman B.F., Mertens B.,
RA Logan-Hendrey L.L.;
RT "Cloning of a cDNA encoding bovine erythropoietin and analysis of its
RT transcription in selected tissues.";
RL Gene 171:275-280(1996).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASSES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC
DR EMBL; L41354; AAB41268.1; -.
DR EMBL; U44762; AAA86653.1; -.
DR HSSP; P01588; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropn.
DR Pfam; PF00758; EPO_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO; 1.
DR Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 25
FT CHAIN 1 192 ERYTHROPOIETIN.
FT FT 32 187 BY SIMILARITY.
FT DISULFID 58 BY SIMILARITY.
FT CARBOHYD 49 49 BY SIMILARITY.
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 192 AA; 21075 MW; DDC419022EF7B483A CRC64;

Query Match 81.9%; Score 692.5; DB 1; Length 192;
Best Local Similarity 83.1%; Pred. No. 2.8e-61;
Matches 138; Conservative 8; Mismatches 19; Indels 1; Gaps 1;

Oy 1 APPRLICDSRYLERYLTAKAEANTTTTCAEHCSTINENTYVDTKYNFAMKRMVEYGOOA 60
Db 26 APARLIGDSRYLERYLTAKAEANTTTTCAEHCSTINENTYVDTKYNFAMKRMVEYGOOA 85
Oy 61 VEWVQGSJLSLSEALLQALQANSSQSPESJQLHDKAVSGJRSJFSLRLVJGAQKELMS 120

```

[illegible]

```

DR EMBL: AF312033; AA28825.1; -.
DR PIR: A24901; A24901.
DR PIR: A24902; A24902.
DR HSSP: P01588; 1CM4.
DR MGD: MGI:95407; EPO.
DR InterPro: IPR001323; EPO_TPO.
DR InterPro: IPR003013; Erythropn.
DR Pfam: PF00758; EPO_TPO.1
DR PRINTS: PR00272; ERYTHROPIN.
DR PROSITE: PS00817; EPO_TPO.1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 26
FT CHAIN 1 26
FT DISULFD 27 192
FT CARBOHYD 33 187
FT CARBOHYD 50 50
FT CARBOHYD 64 64
FT CARBOHYD 109 109
SQ SEQUENCE 192 AA; 21365 MW; 65F94E214E0DEF2E CRC64;

Query Match
Best Local Similarity 81.4%; Score 689; DB 1; Length 192;
Matches 132; Conservative 14; Mismatches 19; Indels 0; Gaps 0;

Oy 1 APPRLICDSRVLEKYLEKKEKENTTGCACGCSINENTTVPDVTYVNFYAKRMVEVGQA 60
Db 27 APPRLICDSRVLEKYLEKKEKENTTGCACGCSINENTTVPDVTYVNFYAKRMVEVEEA 86
Oy 61 VEWOGIALLSFAYRGQALLVNSQPEPLQLHVDKAVSGRSTLTTLRALGAKKEATS 120
Db 87 LEVWOGIALLSFAYRGQALLVNSQPEPLQLHVDKAVSGRSTLTTLRALGAKKEATLS 146

Oy 121 PPDAA-SAAPLTITADPFKLFKRYVSNFLRGKLTLYTGECRTGD 165
Db 147 PPDTPPAPLRLTYDTECKLFKRYVANFLRGKLTLYTGECVCRGD 191

RESULT 8
EPO-SHEEP STANDARD; PRT; 194 AA.
AC P33709; Q28572;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA MEDLINE=93372347; PubMed=8364201;
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czeisuniak J., Goodman M., Bunn H.F.;
RA "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.

```

```

CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z24681; CA80848.1; -.
DR EMBL: L10610; AA31518.1; -.
DR HSSP: P01588; 1CM4.
DR InterPro: IPR001323; EPO_TPO.
DR InterPro: IPR003013; Erythropn.
DR Pfam: PF00758; EPO_TPO.1
DR PRINTS: PR00272; ERYTHROPIN.
DR PROSITE: PS00817; EPO_TPO.1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 26
FT CHAIN 1 26
FT DISULFD 27 192
FT CARBOHYD 34 189
FT CARBOHYD 51 51
FT CARBOHYD 65 65
FT CARBOHYD 110 110
FT CARBOHYD 116 116
FT CONFLICT 108 108
SQ SEQUENCE 194 AA; 21335 MW; C025AAB0528131A9 CRC64;

Query Match
Best Local Similarity 81.9%; Score 685.5; DB 1; Length 194;
Matches 136; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

Oy 1 APPRLICDSRVLEKYLEKKEKENTTGCACGCSINENTTVPDVTYVNFYAKRMVEVGQA 60
Db 28 APPRLICDSRVLEKYLEKKEKENTTGCACGCSINENTTVPDVTYVNFYAKRMVEVGQA 87
Oy 61 VEWOGIALLSFAYRGQALLVNSQPEPLQLHVDKAVSGRSTLTTLRALGAKKEATS 120
Db 88 LEVWOGIALLSFAYRGQALLVNSQPEPLQLHVDKAVSGRSTLTTLRALGAKKEATP 147
Oy 121 PPDAA-SAAPLTITADPFKLFKRYVSNFLRGKLTLYTGECRTGD 165
Db 148 LPDAPSAAPLRIPTVDALSKFRYVSNFLRGKLTLYTGECVCRGD 193

RESULT 9
EPO-PIG STANDARD; PRT; 190 AA.
AC P49157;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor (Fragment).
GN EPO.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA MEDLINE=93372347; PubMed=8364201;
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czeisuniak J., Goodman M., Bunn H.F.;
RA "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A

```

```

CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC      -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL: L10607; AAA31029.1; -.
DR      HSSP: P01588; ICN4.
DR      InterPro: IPR001323; EPO_TPO.
DR      Pfam: PF00758; EPO_TPO; 1.
DR      PROSITE: PS00817; EPO_TPO; 1.
KW      Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT      NON_TER      1
FT      SIGNAL      <1 22
FT      CHAIN      23 190
FT      DISULFID      29 185
FT      DISULFID      51 55
FT      CARBOHYD      46 46
FT      CARBOHYD      60 60
FT      CARBOHYD      105 105
FT      CARBOHYD      168 168
SQ      SEQUENCE      190 AA; 20888 MW; A75BDCCE5077E2A CRC64;

Query Match
Best Local Similarity 80.1%; Score 678; DB 1; Length 190;
Matches 137; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

OY      1 APPRLICDSRVLEERYLLEAKAEANTTTCGAEHCSLNENITVPDTKYNFYAMKMEYGOQA 60
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      23 APPRLICDSRVLEERYLLEAKAEANTTTCGAEHCSLNENITVPDTKYNFYAMKMEYGOQA 82
OY      61 VEWOGALLSEAVLNGQALLVNSQWPPELQIHDVKAVSGLSLTLRALGAQKEAIS 120
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      83 MEWOGALLSEAVLNGQALLVNSQWPPELQIHDVKAVSGLSLTLRALGAQKEAIS 142
OY      121 PPDA--ASAAPLRTTADTFKRLFRYNSNPLRGKLTLYTGECRCRD 165
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      143 LEPDASPSPATPLRTAVDTLCKLFRYNSNPLRGKLTLYTGECRCRD 189

RESULT 10
EPO_CANFA
ID      EPO_CANFA      STANDARD:      PRT:      175 AA.
AC      P33707;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Erythropoietin precursor (Fragment).
GN      EPO.
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A.,
RX      MEDLINE=93372347; PubMed=8364201;
RA      Wen D., Bolssel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA      Czelusniak J., Goodman M., Bunn H.F.;
RT      "Erythropoietin structure-function relationships: high degree of
RT      sequence homology among mammals.";
RL      Blood 82:1507-1516(1993).
CC      -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC      REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC      -1- SUBCELLULAR LOCATION: Secreted.

```

```

CC      -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC      -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL: L13027; AAA30842.1; -.
DR      HSSP: P01588; ICN4.
DR      InterPro: IPR001323; EPO_TPO.
DR      Pfam: PF00758; EPO_TPO; 1.
DR      PROSITE: PS00817; EPO_TPO; 1.
KW      Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT      NON_TER      1
FT      SIGNAL      <1 22
FT      CHAIN      23 175
FT      DISULFID      29 >175
FT      DISULFID      51 55
FT      CARBOHYD      46 46
FT      CARBOHYD      60 60
FT      CARBOHYD      105 105
FT      CARBOHYD      175 175
SQ      SEQUENCE      175 AA; 19193 MW; B504F8DB667B6F4 CRC64;

Query Match
Best Local Similarity 81.0%; Score 638; DB 1; Length 175;
Matches 124; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY      1 APPRLICDSRVLEERYLLEAKAEANTTTCGAEHCSLNENITVPDTKYNFYAMKMEYGOQA 60
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      23 APPRLICDSRVLEERYLLEAKAEANTTTCGAEHCSLNENITVPDTKYNFYAMKMEYGOQA 82
OY      61 VEWOGALLSEAVLNGQALLVNSQWPPELQIHDVKAVSGLSLTLRALGAQKEAIS 120
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      83 LEWOGALLSEAVLNGQALLVNSQWPPELQIHDVKAVSGLSLTLRALGAQKEAIS 142
OY      121 PPDAASAAPLRTTADTFKRLFRYNSNPLRGKLT 153
      |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB      143 LPEEASAPLRTTADTFKRLFRYNSNPLRGKLT 175

RESULT 11
TPO_CANFA
ID      TPO_CANFA      STANDARD:      PRT:      352 AA.
AC      P42705;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE      (C-MPL ligand) (ML) (Megakaryocyte growth and development factor)
GN      (MGDF).
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 24-44.
RX      TISSUE-Kidney;
RX      MEDLINE=94491201; PubMed=8020099;
RA      Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
RA      Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
RA      Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Werewether L.A.,
RA      Clogson C., Hsu E., Hukom M.M., Hornkohl A., Choi E., Pangelinan M.,
RA      Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
RA      Shuter J., Chute H., Basu R., Selander L., Trollinger D., Slew L.,
RA      Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA      Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,

```

RA Pacificci R., Ponting I., Saito C., Wen D., Yung Y.P., Lin H.,  
RA Bosseman R.A.,  
RT "Identification and cloning of a megakaryocyte growth and development  
RT factor that is a ligand for the cytokine receptor Mpl."  
RL Cell 77:1117-1117(1994).  
CC -1 FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION  
CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR  
CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT  
CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.  
CC -1 SUBCELLULAR LOCATION: Secreted.  
CC -1 DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-  
CC -1-TERMINAL. BELONGS TO THE EPO/ TPO FAMILY.  
CC Interpro: IP003028; EPO\_TPO.  
DR Interpro: IP003028; Erythropoietin.  
DR Pfam: PF00758; EPO\_TPO.1  
DR PRINTS: PS00485; THROMBOPO.  
DR PROSITE: PS00817; EPO\_TPO.1  
KW Cytokine; Glycoprotein; Hormone; Signal.  
FT SIGNAL 23  
FT CHAIN 1 352 THROMBOPOIETIN.  
FT DISULEID 28 172 POTENTIAL.  
FT DISULEID 50 106 POTENTIAL.  
FT CARBOHYD 185 185 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 197 197 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 206 206 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 234 234 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 332 332 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 335 335 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 347 347 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 352 AA; 37641 MW; 024F3Ba1B061FBDB8 CRC64;  
  
Query Match 12.9%; Score 109; DB 1; Length 352;  
Best Local Similarity 24.0%; Pred. No. 0.0025;  
Matches 40; Conservative 23; Mismatches 62; Indels 42; Gaps 5;  
  
OY 1 APRRLCDSEVLYERYLLEAKENITGCAEHCSLENTIVDPDKVNFYAKRMEVGOA 60  
DB 24 APP--ACDPRLNKMRIDSHVLSKRSQCPDIYFSTPLVLAHVDFSLSEMTOKRQTA 81  
OY 61 VEWQGLALLSEVLY--RQCALLVNNSQWPEQLQHVAKVAGSLNLTLLFALGAQKEA 118  
DB 82 QDVWGAVALLLDGVLAARGLT-----GPECLSLDQLQSGQVRL 120  
OY 119 I-----SPDDAASAPLRTTADTFEKKRLFRVYSNFIKGLK 154  
DB 121 ILGALQGLGLGTQLDPQGG-----RTTHRPDMALFLSPQQLRKGVR 161  
  
RESULT 12  
TPO\_HUMAN STANDARD; PRT; 353 AA.  
AC P40225; Q13020; Q15790; Q15791; Q15792;  
MC 01-FEB-1993 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Thrombopoietin precursor (Megakaryocyte colony stimulating factor)  
DE (Mpl) (Megakaryocyte leukemia virus oncogene ligand) (C-mpl ligand)  
GN THPO (Megakaryocyte growth and development factor) (MGDF).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Carnivora; Hominoidea; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Petal liver.  
RX MEDLINE=94261202; Pubmed=8202154;  
de Sauvage F.J., Hess P.E., Spencer S.D., Mallow B.E., Gurney A.L.,  
RA Spencer S.A., Darbonne W.C., Henzel W.J., Wong S.C., Huang W.-J.,  
RA Oles K.J., Hultgren B., Solberg L.A., Jr., Goeddel D.V., Watson D.L.,  
RT "Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-mpl  
ligand".

RL Nature 369:533-538(1994).  
 [2] SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE-Fetal Liver.  
 RC MEDLINE=94291201; PubMed=80200099;  
 RA Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.-S., Martin F.,  
 RA Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,  
 RA Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Mesewether L.A.,  
 RA Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangelinan M.,  
 RA Sun Y., Mar V., Meich J., Simonet L., Jacobsen F., Xie C.,  
 RA Shutter J., Chute H., Basu R., Seldner L., Trillingier D., Siou L.,  
 RA Gacela D., Ritali G., Elliott G., Izumi R., Covey T., Crouse J.,  
 RA Pacifici R., Xu W., del Castillo J., Bron J., Cole S., Hu M.C.-T.,  
 RA Bosselman R.A.;  
 RT "Identification and cloning of a megakaryocyte growth and development  
 RT factor that is a ligand for the cytokine receptor MPL.";  
 RL Cell 77:1117-1124(1994).  
 [3] SEQUENCE FROM N.A. (ISOFORM 1).  
 RP MEDLINE=95108091; PubMed=7800916;  
 RA Foster D.C., Sprecher C.A., Grant F.J., Kirmer J.M., Kuipfer J.L.,  
 RA Holly R.D., Whitmore T.E., Helgel M.D., Bell L.A.N., Ching A.F.,  
 RA McGuire V., Hart C., O'hara P.J., Lok S.;  
 RT "Human thrombopoietin: gene structure, cDNA sequence, expression, and  
 RT chromosomal localization.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:13023-13027(1994).  
 [4] SEQUENCE FROM N.A. (ISOFORM 1).  
 RP MEDLINE=95010765; PubMed=7926023;  
 RA Saito T., Akahori H., Seki N., Hori T.-A., Ogami K., Kawamura K.,  
 RA Miyazaki H.;  
 RT "Molecular cloning and chromosomal localization of the human  
 RT thrombopoietin gene.";  
 RL FEBS Lett. 353:57-61(1994).  
 [5] SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RP MEDLINE=95152076; PubMed=7849319;  
 RA Gurney A.L., Kuang W.-J., Xie M.-H., Mallory B.E., Eaton D.L.,  
 RA de Sauvage F.J.;  
 RT "Genomic structure, chromosomal localization, and conserved  
 RT alternative splice forms of thrombopoietin.";  
 RL Blood 85:981-988(1995).  
 [6] SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE=Liver.  
 RC MEDLINE=96015174; PubMed=8537317;  
 RA Kato T., Ogami K., Shimada Y., Iwanatsu A., Sohma Y., Akahori H.,  
 RA Horie K., Kokubo A., Kudo Y., Meeda E., Kobayashi K., Ohashi H.,  
 RA Horzai F., Inoue H., Kawamura K., Miyazaki K.;  
 RT "Purification and characterization of thrombopoietin.";  
 RL J. Biochem. 118:229-236(1995).  
 [7] SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE=Placenta.  
 RC MEDLINE=95122483; PubMed=7822271;  
 RA Chang M., McIninch J., Basu R., Shutter J., Hsu R., Perkins C., Mar V.,  
 RA Suggs S., Weichert A., Li L., Bartley T., Hunt P., Martin F.,  
 RA Samal B., Bogenberger J.;  
 RT "Cloning and characterization of the human megakaryocyte growth and  
 RT development factor (MDF) gene.";  
 RL J. Biol. Chem. 270:511-514(1995).  
 [8] SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).  
 RP Im S.H., Lee W.-S., Chung K.H.;  
 RT "Cloning and sequencing of human thrombopoietin.";  
 RL Submitted (MAY-1996) to the EMBL/Genbank/DBD databases.  
 [9] FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION  
 CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR  
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT  
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.  
 CC [9] SUBCELLULAR LOCATION: Secreted.  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2/TPo-2 AND

```

RESULT 13
POLG_HCVJ8      STANDARD;      PRT;      3033 AA.
AC      P26661;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Genome polypeptide [contains: Capsid protein C (Core protein) (P22);
DE      Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
DE      (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
DE      (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirus)
DE      (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
DE      NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
DE      NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
OS      Hepatitis C virus (isolate HC-J8) (HCV).
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Hepacivirus.
OX      NCBI_Taxid=11115;
RN      1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=92230323; PubMed=1314459;
RA      Okamoto H., Kuri K., Okada S.-I., Yamamoto K., Iizuka H., Tanaka T.,
RA      Fukuda S., Tsuda F., Mishiro S.;
RT      "Full-length sequence of a hepatitis C virus genome having poor
RT      homology to reported isolates: comparative study of four distinct
RT      genotypes."
RL      Virology 188:331-341(1992).
RC      -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC      HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC      NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC      -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC      precursor polypeptide, commonly with Asp or Glu in the P6
CC      position, Cys or Thr in P1 and Ser or Ala in P1'.
CC      -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC      (RNA)(N).
CC      -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC      LIPIDPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS,
CC      PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC      PROTEIN C AND RNA.
CC      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D10988; BAA01761.1; -.
DR      PIR; A40250; GNMVJ8.
DR      HSSP; P27958; 1HEI.
DR      MEROPS; S29_001; -.
DR      MEROPS; U39_001; -.
DR      InterPro; IPR001410; DEAD.
DR      InterPro; IPR002531; HCV_NS1.
DR      InterPro; IPR002518; HCV_NS2.
DR      InterPro; IPR004109; HCV_NS3.
DR      InterPro; IPR000745; HCV_NS4a.
DR      InterPro; IPR001490; HCV_NS4b.
DR      InterPro; IPR002868; HCV_NS5a.
DR      InterPro; IPR002166; HCV_RdRP.
DR      InterPro; IPR002521; HCV_capsid.
DR      InterPro; IPR002519; HCV_core.
DR      Pfam; PF00998; HCV_RdRP; 1.
DR      Pfam; PF01001; HCV_NS4b; 1.
DR      Pfam; PF01006; HCV_NS4a; 1.
DR      Pfam; PF01506; HCV_NS5a; 1.

```





| Db         | 408  | LEBQJQEVRRK                           | -----CAEQAQJSSLSKAEITTSQESQISSYEBBLKARELSNQDET | 461                        |
|------------|--|---------------------------------------|--|----------------------------|
| Qy         | 68   | ALLSEAVLRGQALLVNSQPEPIQLAHVKAIVSGLRSL | 105  |                            |
| Db         | 462  | AQLESVESGKAQL                         | -----EPILOQHJQESQOEISSM                        | 492                        |
| RESULT 15  |  |                                       |  |                            |
| ATCS_SYNP7 |  | STANDARD:                             | FRF:   | 747 AA.                    |
| ID         | ATCS_SYNP7   |                                       |  |                            |
| AC         | P37279.  |                                       |  |                            |
| DT         | 01-OCT-1994 (Rel. 30, Created)   |                                       |  |                            |
| DT         | 01-OCT-1994 (Rel. 30, Last sequence update)  |                                       |  |                            |
| DT         | 15-JUN-2002 (Rel. 41, Last annotation update)  |                                       |  |                            |
| DE         | Cation-transporting ATPase pacs (EC 3.6.3.-).  |                                       |  |                            |
| GN         |  |                                       |  |                            |
| OS         | Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).   |                                       |  |                            |
| OC         | Bacteria; Cyanobacteria; Chroococcales; Synechococcus.   |                                       |  |                            |
| OX         | NCBI_TaxID=1140;   |                                       |  |                            |
| RN         | [1]  |                                       |  |                            |
| RP         | SEQUENCE FROM N.A.   |                                       |  |                            |
| RX         | MEDLINE=93380581; PubMed=8370468;  |                                       |  |                            |
| RA         | Kanamaru K., Kashiwagi S., Mizuno T.;  |                                       |  |                            |
| RT         | "The cyanobacterium, Synechococcus sp. PCC7942, possesses two  |                                       |  |                            |
| RT         | distinct genes encoding cation-transporting P-type ATPases.";  |                                       |  |                            |
| RL         | FEBS Lett. 330:99-104(1993).   |                                       |  |                            |
| CC         | -1- FUNCTION: MAY PLAY A ROLE IN THE OSMOTIC ADAPTATION.   |                                       |  |                            |
| CC         | -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.   |                                       |  |                            |
| CC         | -1- SUBCELLULAR LOCATION: Integral membrane protein.   |                                       |  |                            |
| CC         | -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY   |                                       |  |                            |
| CC         | (EL-E2 ATPASES). SUBFAMILY 1B.   |                                       |  |                            |
| CC         | -1- SIMILARITY: CONTAINS 1 HMA DOMAIN.   |                                       |  |                            |
| CC         | -----  |                                       |  |                            |
| CC         | This SWISS-PROT entry is copyright. It is produced through a collaboration   |                                       |  |                            |
| CC         | between the Swiss Institute of Bioinformatics and the EMBL outstation  |                                       |  |                            |
| CC         | the European Bioinformatics Institute. There are no restrictions on its  |                                       |  |                            |
| CC         | use by non-profit institutions as long as its content is in no way   |                                       |  |                            |
| CC         | modified and this statement is not removed. Usage by and for commercial  |                                       |  |                            |
| CC         | entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |                                       |  |                            |
| CC         | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).  |                                       |  |                            |
| CC         | -----  |                                       |  |                            |
| DR         | EMBL: D16437; BAA03907.1; -.   |                                       |  |                            |
| DR         | PIR: S36741; S36741.   |                                       |  |                            |
| DR         | HSSP: Q04656; IAM0.  |                                       |  |                            |
| DR         | InterPro: IPR001757; ATPase_E1-E2.   |                                       |  |                            |
| DR         | InterPro: IPR001934; HeavyMe_transpt.  |                                       |  |                            |
| DR         | InterPro: IPR001454; Hlgase/hydrilase.   |                                       |  |                            |
| DR         | Pfam: PF001122; E1-E2_ATPase; 1.   |                                       |  |                            |
| DR         | Pfam: PF00403; HMA; 1.   |                                       |  |                            |
| DR         | Pfam: PF00702; Hydrolase; 1.   |                                       |  |                            |
| DR         | PROSITE: PS00154; ATPase_E1_E2; 1.   |                                       |  |                            |
| DR         | PROSITE: PS01047; HMA_1; 1.  |                                       |  |                            |
| DR         | Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;   |                                       |  |                            |
| KW         | Metal-binding.   |                                       |  |                            |
| FT         | DOMAIN 1   | 101                                   |  | CYTOPLASMIC (POTENTIAL).   |
| FT         | TRANSMEM   | 102                                   | 122  | POTENTIAL.                 |
| FT         | DOMAIN   | 123                                   | 132  | EXTRACELLULAR (POTENTIAL). |
| FT         | TRANSMEM   | 133                                   | 151  | POTENTIAL.                 |
| FT         | DOMAIN   | 152                                   | 158  | CYTOPLASMIC (POTENTIAL).   |
| FT         | TRANSMEM   | 159                                   | 179  | POTENTIAL.                 |
| FT         | DOMAIN   | 180                                   | 199  | EXTRACELLULAR (POTENTIAL). |
| FT         | TRANSMEM   | 200                                   | 220  | POTENTIAL.                 |
| FT         | DOMAIN   | 221                                   | 348  | CYTOPLASMIC (POTENTIAL).   |
| FT         | TRANSMEM   | 349                                   | 371  | POTENTIAL.                 |
| FT         | DOMAIN   | 372                                   | 378  | EXTRACELLULAR (POTENTIAL). |
| FT         | TRANSMEM   | 379                                   | 396  | POTENTIAL.                 |
| FT         | DOMAIN   | 397                                   | 688  | CYTOPLASMIC (POTENTIAL).   |
| FT         | TRANSMEM   | 689                                   | 708  | POTENTIAL.                 |
| FT         | DOMAIN   | 709                                   | 720  | EXTRACELLULAR (POTENTIAL). |
| FT         | TRANSMEM   | 721                                   | 739  | POTENTIAL.                 |
| FT         | DOMAIN   | 740                                   | 747  | CYTOPLASMIC (POTENTIAL).   |
| FT         | DOMAIN   | 740                                   | 770  | HMA.                       |

| FT | MOD_RES  | 434     | 434       | PHOSPHORYLATION (BY SIMILARITY) |
|----|----------|---------|-----------|---------------------------------|
| FT | METAL    | 14      | 14        | POTENTIAL.                      |
| FT | METAL    | 17      | 17        | POTENTIAL.                      |
| FT | METAL    | 634     | 634       | MAGNESIUM (BY SIMILARITY).      |
| FT | METAL    | 638     | 638       | MAGNESIUM (BY SIMILARITY).      |
| SQ | SEQUENCE | 747 AA: | 79732 MW: | QAF18B92EB85CEE CRC64;          |

|                       |       |              |     |            |    |        |     |
|-----------------------|-------|--------------|-----|------------|----|--------|-----|
| Query Match           | 9.28  | Score        | 78  | DB         | 1  | Length | 747 |
| Best Local Similarity | 25.48 | Pred. NO.    | 7.3 |            |    |        |     |
| Matches               | 33    | Conservative | 22  | Mismatches | 45 | Indels | 30  |
|                       |       |              |     |            |    | Gaps   | 5   |

```

QY 17 LEAKAEVITTTGCAHCSLNNITVPTQKVFYMKRMVEWQOAV---EYWGGLATLSEA 73
Db 233 LQAKTARVLRQG-----QELPLITEVEDEWVRVPEKEKVPVDEVIDGSTVDES 284
QY 74 VLKRGALLVNSQPEWELQLHVKAVSLR-----SLTTLRLALG---AQKEAISPDP 123
Db 285 MYTGSL-----PVQKQGVGEVIGATLNTKGTSLTIRATRVGRETFILAQIVQLVQQA 335
QY 124 AASAPLRTI 133
Db 336 QASKAPIQLR 345

```

Search completed: January 2, 2003, 15:11:51  
Job time : 10.4743 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:09:40 : Search time 26.4199 Seconds  
(without alignments) updates/sec  
1286.823 Million cell

Title: US-09-853-731-1

Perfect score: 846  
Sequence: 1 APPRLICDSRVLEERYLLLEAK.....SNFLGKLTLYTGACRTGD 165

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 680.5 | 80.4        | 195    | 6 O9GKA3  | O9gka3 oryctolagus |
| 2          | 680.5 | 80.4        | 195    | 6 O9GKA2  | O9gka2 oryctolagus |
| 3          | 678   | 80.1        | 194    | 6 O9MYM8  | O9mym8 sus scrofa  |
| 4          | 188   | 22.2        | 50     | 11 O9OV40 | O9ov40 rattus sp.  |
| 5          | 88    | 10.4        | 323    | 16 O8ZDC8 | O8zdc8 yersinia pe |
| 6          | 87.5  | 10.3        | 346    | 16 O8ZKZ4 | O8zkz4 salmonella  |
| 7          | 87.5  | 10.3        | 346    | 16 O8Z2M5 | O8z2m5 salmonella  |
| 8          | 85    | 10.0        | 3722   | 2 P94873  | P94873 lysobacter  |
| 9          | 83    | 9.8         | 296    | 16 O8ZAY4 | O8zay4 yersinia pe |
| 10         | 83    | 9.8         | 339    | 16 O9HZM7 | O9hzm7 pseudomonas |
| 11         | 82.5  | 9.8         | 3033   | 12 O9DHD6 | O9dhd6 hepatitis c |
| 12         | 82    | 9.7         | 815    | 10 O9FK91 | O9fk91 arabidopsis |
| 13         | 79.5  | 9.4         | 1829   | 16 O86808 | O86808 streptomyce |
| 14         | 79.5  | 9.4         | 3033   | 12 O9IZAI | O9izai hepatitis c |
| 15         | 79    | 9.3         | 480    | 10 O43380 | O43380 avena sativ |
| 16         | 79    | 9.3         | 2364   | 5 Q22896  | Q22896 caenorhabdt |

|    |      |     |      |           |                     |
|----|------|-----|------|-----------|---------------------|
| 17 | 78.5 | 9.3 | 813  | 16 O8Z9B9 | O8z9b9 salmonella   |
| 18 | 78.5 | 9.3 | 3019 | 12 O92529 | O92529 hepatitis c  |
| 19 | 78   | 9.2 | 348  | 16 O86751 | O86751 streptomyce  |
| 20 | 78   | 9.2 | 378  | 17 O8ZK3  | O8zk3 pyrobaculum   |
| 21 | 78   | 9.2 | 455  | 16 O8UB09 | O8ub09 agrobacteri  |
| 22 | 78   | 9.2 | 548  | 2 O93148  | O93148 buchnera ap  |
| 23 | 78   | 9.2 | 548  | 2 O93N35  | O93n35 buchnera ap  |
| 24 | 77.5 | 9.2 | 242  | 16 O8YV76 | O8yv76 anabaena sp  |
| 25 | 77.5 | 9.2 | 310  | 16 O91033 | O91033 streptomyce  |
| 26 | 77.5 | 9.2 | 451  | 16 O8XSE8 | O8xse8 raietonia s  |
| 27 | 77.5 | 9.2 | 1003 | 10 O9FIR4 | O9fir4 arabidopsis  |
| 28 | 77.5 | 9.2 | 2556 | 11 O91XW2 | O91xw2 mus musculu  |
| 29 | 77   | 9.1 | 379  | 2 O9L9F5  | O9l9f5 streptomyce  |
| 30 | 77   | 9.1 | 383  | 16 O9ZOV7 | O9zov7 rhizobium m  |
| 31 | 77   | 9.1 | 447  | 2 P72270  | P72270 rhodococcus  |
| 32 | 77   | 9.1 | 451  | 16 P74054 | P74054 synecocyst   |
| 33 | 76.5 | 9.0 | 154  | 16 O9PGAT | O9pgat xylella tas  |
| 34 | 76.5 | 9.0 | 425  | 16 O8YF20 | O8yf20 bruceella me |
| 35 | 76.5 | 9.0 | 619  | 3 O96VB7  | O96vb7 botrytis cl  |
| 36 | 76.5 | 9.0 | 1315 | 3 O8X215  | O8x215 botrytis cl  |
| 37 | 76.5 | 9.0 | 1315 | 3 O8X1E7  | O8x1e7 botrytis cl  |
| 38 | 75.5 | 8.9 | 637  | 16 O5413  | O5413 synecocyst    |
| 39 | 75.5 | 8.9 | 871  | 5 O9NB04  | O9nb04 drosophila   |
| 40 | 75   | 8.9 | 397  | 8 O9TNG5  | O9tng5 coriaria te  |
| 41 | 75   | 8.9 | 308  | 10 O94141 | O94141 oryza sativ  |
| 42 | 75   | 8.9 | 548  | 2 O93N34  | O93n34 buchnera ap  |
| 43 | 75   | 8.9 | 554  | 2 O9RPH5  | O9rph5 mycobacteri  |
| 44 | 74.5 | 8.8 | 361  | 6 O9T5W6  | O9tsw6 sus scrofa   |
| 45 | 74.5 | 8.8 | 389  | 2 O9AN18  | O9an18 bradyrhizob  |

## ALIGNMENTS

|                       |  |                             |           |             |  |
|-----------------------|--|-----------------------------|-----------|-------------|--|
| RESULT 1              |  |                             |           |             |  |
| O9GKA3                | O9GKA3   | PRELIMINARY;                | PRT;      | 195 AA.     |  |
| AC                    | O9GKA3:  |                             |           |             |  |
| DT                    | 01-MAR-2001 (TREMBLE)  | 16, Created                 |           |             |  |
| DT                    | 01-MAR-2001 (TREMBLE)  | 16, Last sequence update)   |           |             |  |
| DT                    | 01-JUN-2002 (TREMBLE)  | 21, Last annotation update) |           |             |  |
| DE                    | Erythropoietin.  |                             |           |             |  |
| OS                    | Oryctolagus cuniculus (Rabbit).                                    |                             |           |             |  |
| OC                    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |                             |           |             |  |
| OC                    | Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.            |                             |           |             |  |
| OX                    | NCBI_TaxID=9986;   |                             |           |             |  |
| RN                    | [1]  |                             |           |             |  |
| RP                    | SEQUENCE FROM N.A.   |                             |           |             |  |
| RX                    | MEDLINE-21290682; PubMed-11396976;                                 |                             |           |             |  |
| RA                    | Villalta A., Wu D., Margalith M., Hobart P.;                       |                             |           |             |  |
| RT                    | "Rabbit EPO Gene and cDNA: Expression of Rabbit EPO after          |                             |           |             |  |
| RT                    | Intramuscular Injection of pDNA."                                  |                             |           |             |  |
| RL                    | Biochem. Biophys. Res. Commun. 284:823-827(2001).                  |                             |           |             |  |
| DR                    | EMBL: AF290943; AAC36961.1; -                                      |                             |           |             |  |
| DR                    | HSSP: P01588; ICN4.  |                             |           |             |  |
| DR                    | InterPro: IPR001323; EPO.TPO.                                      |                             |           |             |  |
| DR                    | InterPro: IPR003013; Erythropo.                                    |                             |           |             |  |
| DR                    | PIRfam: PF00758; EPO.TPO.1.  |                             |           |             |  |
| DR                    | PRINTS: PR00272; ERYTHROPTN.                                       |                             |           |             |  |
| DR                    | PROSITE: PS00817; EPO.TPO.1.                                       |                             |           |             |  |
| SQ                    | SEQUENCE 195 AA; 21053 MW; 0999DA7B52713F3 CRC64;                  |                             |           |             |  |
| Query Match           | 80.4%;   | Score 680.5;                | DB 6;     | Length 195; |  |
| Best Local Similarity | 81.3%;   | Pred. No. 1.4e-60;          |           |             |  |
| Matches 195;          | Conservative 12;   | Mismatches 18;              | Indels 1; | Gaps 1;     |  |
| OY                    | 1 APPRLICDSRVLEERYLLLEAKAEENITGCAEHGSLNINITYPPDKVNFYAKKRMVEGOA 60  |                             |           |             |  |
| DB                    | 29 APPRLICDSRVLEERYLLLEAKAEENITGCAEHGSLNINITYPPDKVNFYAKKRMVEGOA 88 |                             |           |             |  |
| OY                    | 61 VEWOGTALLSEAVLRGQALLVNSQSPWEPQLHVDRAVSGRLSTLTLLRALGAKGKAIS 120  |                             |           |             |  |
| DB                    | 89 VEWOGTALLSEAVLRGQALLVNSQSPWEPQLHVDRAVSGRLSTLTLLRALGAKGKAIS 148  |                             |           |             |  |

QY 121 PPDA--SAAPLRTTADTFKRLFRVYSNFLRGKLTLYTGACRTGD 165  
 DB 149 PPEAASSAAPLRTVADTLCKLFRYSNFLRGKLTLYTGACRRGD 194

## RESULT 2

Q9GKA2 PRELIMINARY: PRT; 195 AA.

AC 09GKA2: 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 OS Erythropoietin  
 OC Erythrocytes cuniculus (Rabbit)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA MEDLINE=21290682; PubMed=11396976;  
 RX Vilalta A., Wu D., Margalit M., Hobart P.,  
 RT "Rabbit EPO Gene and cDNA: Expression of Rabbit EPO after  
 RT Intramuscular Injection of pDNA"  
 RL Biochem. Biophys. Res. Commun. 284:823-827(2001).  
 DR HSSP; P01588; 1CN4.  
 DR InterPro; IPR001323; EPO\_TPO.  
 DR Pfam; PF00758; EPO\_TPO; 1.  
 DR PRINTS; PR00272; ERYTHROPTN.  
 DR PROSITE; PS00817; EPO\_TPO; 1.  
 SQ SEQUENCE 195 AA; 21025 MW; 1F1DC7F403A303BC CRC64;

Query Match 80.1%; Score 680.5; DB 6; Length 195;  
 Best Local Similarity 81.4%; Pred. No. 1,4e-60;  
 Matches 133; Conservative 12; Mismatches 18; Indels 1; Gaps 1;

QY 1 APPRLIDSRVLEKAEENITTCAGHCSLNTITVPPDKVNFYAMKMEVGOOA 60  
 DB 29 AAPRLIDSRVLEKAEENITTCAGHCSLNTITVPPDKVNFYAMKMEVGOOA 88  
 QY 61 VEWOGIALISEAVLRGOALIVNSQWPEPLQIHDVKAVSGSLTTLRALGAOKEAIS 120  
 DB 89 VEWOGIALISEAVLRGOALIVNSQWPEPLQIHDVKAVSGSLTTLRALGAOKEAIS 148  
 QY 121 PPDA--SAAPLRTTADTFKRLFRVYSNFLRGKLTLYTGACRTGD 165  
 DB 149 PPEAASSAAPLRTVADTLCKLFRYSNFLRGKLTLYTGACRRGD 194

## RESULT 3

Q9MYM8 PRELIMINARY: PRT; 194 AA.

AC 09MYM8: 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 OS Erythropoietin precursor.  
 OC Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=NORWEGIAN LANDRACE; TISSUE=KIDNEY;  
 RA David B. Harbige, J.,  
 RT "The porcine erythropoietin gene: cDNA and genomic sequences and  
 RT expression analysis"  
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249745; CAB96416.1; -.  
 DR EMBL; AJ249746; CAB96417.1; -.

DR HSSP; P01588; 1CN4  
 DR InterPro; IPR001323; EPO\_TPO.  
 DR InterPro; IPR001323; Erythroptn.  
 DR Pfam; PF00758; EPO\_TPO; 1.  
 DR PRINTS; PR00272; ERYTHROPTN.  
 DR PROSITE; PS00817; EPO\_TPO; 1.  
 KW Signal.  
 FT SIGNAL. 1 26 POTENTIAL.  
 FT CHAIN 27 194 ERYTHROPTN.  
 SQ SEQUENCE 194 AA; 21303 MW; 778B1A6F62DEA1C CRC64;

Query Match 80.1%; Score 678; DB 6; Length 194;  
 Best Local Similarity 82.0%; Pred. No. 2.5e-60;  
 Matches 137; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

QY 1 APPRLIDSRVLEKAEENITTCAGHCSLNTITVPPDKVNFYAMKMEVGOOA 60  
 DB 27 AAPRLIDSRVLEKAEENITTCAGHCSLNTITVPPDKVNFYAMKMEVGOOA 86  
 QY 61 VEWOGIALISEAVLRGOALIVNSQWPEPLQIHDVKAVSGSLTTLRALGAOKEAIS 120  
 DB 87 VEWOGIALISEAVLRGOALIVNSQWPEPLQIHDVKAVSGSLTTLRALGAOKEAIP 146  
 QY 121 PPDA--SAAPLRTTADTFKRLFRVYSNFLRGKLTLYTGACRTGD 165  
 DB 147 LPDAPSSAAPLRTVADTLCKLFRYSNFLRGKLTLYTGACRRGD 193

## RESULT 4

Q9GV40 PRELIMINARY: PRT; 50 AA.

AC 09GV40: 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 OS Erythropoietin (fragment).  
 OC Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE.

RA MEDLINE=94115047; PubMed=7764337;  
 RX Okano M., Suga H., Masuda S., Nagao M., Narita H., Ikura K.,  
 SA Sasaki R.,  
 RL Biosci. Biotechnol. Biochem. 57:1882-1885(1993).  
 DR HSSP; P01588; 1EER.  
 DR InterPro; IPR001323; EPO\_TPO.  
 DR Pfam; PF00758; EPO\_TPO; 1.  
 DR PRINTS; PR00272; ERYTHROPTN.  
 SQ SEQUENCE 50 AA; 5587 MW; 70B44A8BFEO16034 CRC64;

Query Match 22.2%; Score 188; DB 11; Length 50;  
 Best Local Similarity 76.0%; Pred. No. 9.9e-12;  
 Matches 35; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 APPRLIDSRVLEKAEENITTCAGHCSLNTITVPPDKVNFYAMKMEVGOOA 50  
 DB 1 APPRLIDSRVLEKAEENITTCAGHCSLNTITVPPDKVNFYAMKMEVGOOA 50

## RESULT 5

Q8ZDC8 PRELIMINARY: PRT; 323 AA.

AC 08ZDC8: 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 OS Ribonucleoside-diphosphate reductase 2 beta chain (EC 1.17.4.1).  
 GN NNDP OR YPO2648.  
 OS Versinia pectis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```

OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parikh H.J., Wren B.W., Thomson N.R., Fitball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
RA Leathem S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds S., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL: AJ414153; CAC92889.1.
DR InterPro: IPR000358; RibonucL_redctse.
DR Pfam: PF00268; ribonuc_red_sm; 1.
DR PROSITE: PS00368; RIBORED_SMALL; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 323 AA; 36823 MW; 87C21F7BB9B7FD2 CRC64;

Query Match 10.4%; Score 88; DB 16; Length 323;
Best Local Similarity 25.2%; Pred. No. 1.3;
Matches 34; Conservative 20; Mismatches 59; Indels 22; Gaps 5;

OY 38 NITVVDITVNFYAKRMVEVGOAVVWOGALISEAVLRGQALLVNSSOPWEPLOLHYD- 96
DB 2 NVKFKITRISALNMKRIE-DDKDELVWN--RLTSNFWLPEKVPISNDIPSWATLTPHEQO 58
OY 97 ---KAVSGLSRLITLRLALGAO---KEAISPDAASAPLRTITDTRFKFRVYSNFLR 150
DB 59 LTRIFRTGLITDITONTLGAALIKDALTTPHEAIFSNISMEVHARSYSTSTL-- 116
OY 151 GKLLKTYGEACRTGD 165
DB 117 -----CLTSPD 121

RESULT 6
O8KX24 PRELIMINARY; PRT; 346 AA.
AC O8KX24;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Periplasmic sensor in multi-component regulatory system with Tors
DE (sensory kinase) and TorR (regulator), regulates tor operon.
GN TORR OR STM3825.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lt2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvanev E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT Lt2.";
RL Nature 413:852-856(2001).
DR EMBL: AE008878; AL22684.1.
DR InterPro: IPR001761; PeriplabP/LacI.
DR Pfam: PF00532; Peripla_BP_like; 1.
KW Complete proteome.
SQ SEQUENCE 346 AA; 38440 MW; E37CAB58E49FD716 CRC64;

Query Match 10.3%; Score 87.5; DB 16; Length 346;
Best Local Similarity 26.7%; Pred. No. 1.5;

```

```

Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

OY 10 RYLERVYLEAKAEVNTTG--CAEHCISLNE--NITVPDTKVFYAKRMVEVGOAVVWQ 65
DB 217 RNLQDEMERHPDANVAVGSAIAEAMGEGRNLTPTIVSFYL-----THQVYR 267
OY 66 GIALISEAVLRGQALLVNSSQ--PWEPLQHYDKAVSGLSLITLRLALGAO--KEAISP 122
DB 268 GLK-----RHIIMALSDQAMQ-----GELAIQSIKIVLOGQVPEINISPP 309
OY 123 -----DAASAPLRTITADTRFKLFRVYSNFLRKLKLYGEEA 160
DB 310 VLILTHNNADSARVRSLSPPGFRPVY-----LYQYTSSEA 344

RESULT 7
O8K2M5 PRELIMINARY; PRT; 346 AA.
AC O8K2M5;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Solute binding receptor protein.
GN STY3952.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parikh H.J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,
RA Krogh A., Larsen T.S., Leathem S., Moule S., O'Gaora P., Barry C.,
RA Quail M., Rutherford K., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL: AL627280; CAD03169.1.
DR InterPro: IPR001761; PeriplabP/LacI.
DR Pfam: PF00532; Peripla_BP_like; 1.
KW Receptor; Complete proteome.
SQ SEQUENCE 346 AA; 38546 MW; F80FB168B8C3A8F CRC64;

Query Match 10.3%; Score 87.5; DB 16; Length 346;
Best Local Similarity 26.7%; Pred. No. 1.5;
Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

OY 10 RYLERVYLEAKAEVNTTG--CAEHCISLNE--NITVPDTKVFYAKRMVEVGOAVVWQ 65
DB 217 RNLQDEMERHPDANVAVGSAIAEAMGEGRNLTPTIVSFYL-----THQVYR 267
OY 66 GIALISEAVLRGQALLVNSSQ--PWEPLQHYDKAVSGLSLITLRLALGAO--KEAISP 122
DB 268 GLK-----RHIIMALSDQAMQ-----GELAIQSIKIVLOGQVPEINISPP 309
OY 123 -----DAASAPLRTITADTRFKLFRVYSNFLRKLKLYGEEA 160
DB 310 VLILTHNNADSARVRSLSPPGFRPVY-----LYQYTSSEA 344

RESULT 8
P94873 PRELIMINARY; PRT; 3722 AA.
AC P94873;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

```

DE Alpha-aminoadipyl-cysteine-lysinase synthetase.  
 GN PCBA.  
 OS Lysobacter lactamgenus.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 RX Lysobacter.  
 OC NCBI\_TaxID=39596;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K90;  
 RX MEDLINE=96275949; PubMed=8737573;  
 RT Kimura H., Miyashita H., Sumino Y.;  
 RT "Organization and expression in *Pseudomonas putida* of the gene cluster  
 RT involved in cephalosporin biosynthesis from *Lysobacter lactamgenus*  
 RT K90.";  
 RT APPL. Microbiol. Biotechnol. 45:490-501(1996).  
 DR EMBL: D50308; BAA0846.1; -.  
 DR HSSP: P14687; 1AMU.  
 DR InterPro: IPR002106; AAcRNA\_LigaseII.  
 DR InterPro: IPR000873; AMP-bind.  
 DR InterPro: IPR001242; Condensatn.  
 DR InterPro: IPR009777; DNA\_Ligase.  
 DR InterPro: IPR003880; Pantine\_attach.  
 DR InterPro: IPR000379; Ser\_ester\_site.  
 DR Pfam: PF00501; AMP-binding; 3.  
 DR Pfam: PF00668; Condensation; 3.  
 DR Pfam: PF00975; Thioesterase; 1.  
 DR PROSITE: PS00179; AA\_TRNA\_LIGASE\_II\_1; UNKNOWN\_1.  
 DR PROSITE: PS00075; ACP\_DOMAIN; 3.  
 DR PROSITE: PS00455; AMP\_BINDING; 1.  
 DR PROSITE: PS00697; DNA\_LIGASE\_A1; UNKNOWN\_2.  
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN\_2.  
 KW Phosphopantetheine.  
 SO SEQUENCE 3722 AA; 411607 MW; 3597B3483463809B CRC64;

Query Match Best Local Similarity 10.0%; Score 85; DB 2; Length 3722;  
 Matches 34; Conservative 18; Mismatches 50; Indels 52; Gaps 5;  
 QY 16 LLEAKEANITGGCAHCSINENITVPDTKVFYAMKREYGOQAVEV---WQGLALLS 71  
 DB 2138 LLEGATACATLP-----PLPATYADFSVWQROQSDORDALPDYWO----- 2178  
 QY 72 EAVLRGQALLVNSQPEPIQLHVDKAVSGLSLTLRLALGAQ-----KEAISPDAAS 126  
 DB 2179 -----RSLAGWQPIQLPLD-----HARPAQFDYLGREIVFDVADAT 2214  
 QY 127 AAPLRITADTFKRLFRVYSNFKLGLKLYTGEA 160  
 DB 2215 CDQLRLVLAQTRTSFESVLLAAVYLLTKAVSGS 2248

RESULT 9  
 ID 08ZAY4 PRELIMINARY; PRT; 296 AA.  
 AC 08ZAY4;  
 DT 01-MAR-2002 (TREMblrel. 20, Created)  
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE putative 2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60).  
 GN yro3648.  
 OS *Yersinia pestis*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Yersinia*.  
 OC NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92; BIOVAR ORIENTALIS;  
 RX MEDLINE=21470413; PubMed=11586360.  
 RA Partridge J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,  
 RA Prentice H.B., Sebaila M., James K.D., Churcher C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;  
 RT "Genome sequence of *Yersinia pestis*, the causative agent of plague".  
 RL Nature 413:523-527(2001).  
 DR EMBL: AJ414158; CAC93117.1; -.  
 DR InterPro: IPR002204; 3hydroxisobut\_dh.  
 DR InterPro: IPR001744; 6PGD.  
 DR Pfam: PF03446; NAD\_binding\_2; 1.  
 DR PROSITE: PS00895; 3-HYDROXYISOBUT\_DH; UNKNOWN\_1.  
 KW Oxidoreductase; Hypothetical protein; Complete proteome.  
 SO SEQUENCE 296 AA; 32140 MW; C883AFCC6868429D CRC64;

Query Match Best Local Similarity 9.8%; Score 83; DB 16; Length 296;  
 Matches 48; Conservative 34; Mismatches 62; Indels 54; Gaps 10;  
 QY 6 ICDSRVLERYLLLEAK-EAENITGGCAHCSINENITVPDTKVFYAMKREYGOQAVEV 63  
 DB 32 LCDADITKVOILTALNAESITVDNA--ASIDLITLWLPNSE--AVEQVLGSDGISG 85  
 QY 64 WQGLALLSEAVLRGQALLVNSQPEPIQLHVDKAVSG-----LRTTADTFKRL 142  
 DB 86 W--VAQLSQAAV--VIDMSSSDPERSRLAILLAWELDYLDAPVSGGVKKNQCTLSI 140  
 QY 102 -----LNSLTTLRALGAQKEAISPDAASAP-----LRTTADTFKRL 142  
 DB 141 LIGGEDRYLKSCTALAMEQDILFVGPAAGSHAAALNNYSATGILATIEALHVAQRF 200  
 QY 143 RYVSNFLRGKLYTGEA 160  
 DB 201 GIEPEVTEVLTNLTGNS 218

RESULT 10  
 ID 09HZM7 PRELIMINARY; PRT; 339 AA.  
 AC 09HZM7;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DE UDP-N-acetylglucosamine 6-phosphate 4-epimerase (EC 5.1.1.11).  
 GN MURB OR PA2977.  
 OS *Pseudomonas aeruginosa*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC *Pseudomonas*.  
 OC NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Phan X.Q., Erwin A.L., Mizoguchi S.D., Warrenner P.,  
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.D., Lagrou M.,  
 RA Garber R.L., Goltry L.A., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
 RA Bardsley L.L., Coulter S.A., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
 RA Smith K.A., Spencer D.H., Wong G.K.S., Kas A., Lardig K., Lim R.M.,  
 RA Reizer J., Sailer M.H., Hancock R.B.W., Lory S., Paulsen I.T.,  
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an  
 RT opportunistic pathogen".  
 RL Nature 406:959-964(2000).  
 DR EMBL: AE004723; AAG06365.1; -.  
 DR HSSP: P08373; 2MBR.  
 DR InterPro: IPR003170; MURB.  
 DR InterPro: IPR001575; OXID\_PAD\_bind.  
 DR Pfam: PF01565; PAD\_binding\_4; 1.  
 DR Pfam: PF02873; MURB\_C; 1; 1.  
 DR PROSITE: PS00430; TONE\_DEPENDENT\_REC\_1; UNKNOWN\_1.  
 KW Complete proteome.  
 SO SEQUENCE 339 AA; 37627 MW; C0C8EF92938FE27 CRC64;  
 Query Match 9.8%; Score 83; DB 16; Length 339;

Best Local Similarity 22.7%; Pred. No. 4.3;  
Matches 41; Conservative 22; Mismatches 48; Indels 70; Gaps 8;

```
OY 31 EHCISLNE-NITVPTKVFYAKRMVEGQAVEWQGLALLSEALRQCALLV----- 82
      ||||| | : | | | | : | | | | : | | | | : | | | | : | | | |
Db 7 EHCISLKPNTGIDVRAFLAHARDE-----ADVREALALARE--RCLPLLVIGGCSNL 58
OY 83 -----NSSQWPPEP-LQLHVDAKAVSGLSRLTLL 109
Db 59 LITRDVEALVLRMASQGRRIYSDAASVLAEEAGEAMDPEYQWSLGRGLALENLSIL- 117
OY 110 RALGAQKEAISPPDAASAPLRTITR-----DTRKLFERYVSNPLRGKLTLYGEACRT 163
Db 118 -----PCTVGAAPMKNIGAYGVELKDVFDLSLAL--DNQDGLREFDQACRF 163
OY 164 G 164
Db 164 G 164
```

## RESULT 11

```
Q9DHD6 PRELIMINARY; PRT; 3033 AA.
ID Q9DHD6
AC Q9DHD6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)].
OS Hepatitis C virus type 2b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31650;
RN [1]
RN SEQUENCE FROM N.A.
RC MEDLINE=21296595; PubMed=11402859;
RA Murakami K., Abe M., Kageyama T., Kamoshita N., Nomoto A.;
RT "Down-regulation of translation driven by Hepatitis C virus internal
RT ribosomal entry site by the 3' untranslated region of RNA.";
RL Arch. Virol. 146:729-741(2001).
CC -1 SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL: AB030907; BAB08107.1; -
DR HSSP: p27958.1A1Y.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR002518; HCV_NS2.
DR InterPro: IPR004109; HCV_NS3.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRp.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; HCV_RdRp; 1.
DR ProDom: PD186062; HCV_NS1; 1.
DR SMART: SM00487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KM Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KM Polypeptide; RNA-directed RNA polymerase; Transmembrane.
FT CHAIN 1 191 CORE PROTEIN.
FT CHAIN 192 383 E1 PROTEIN.
```

```
FT CHAIN 384 750 E2 PROTEIN.
FT CHAIN 751 813 p7 PROTEIN.
FT CHAIN 1030 NS2 PROTEIN.
FT CHAIN 1031 NS3 PROTEINASE/HELICASE.
FT CHAIN 1661 NS4A PROTEIN.
FT CHAIN 1662 1715 NS4B PROTEIN.
FT CHAIN 1716 1976 NS5A PROTEIN.
FT CHAIN 1977 2442 NS5B PROTEIN.
FT CHAIN 2443 3033 NS5B RNA-DEPENDENT RNA POLYMERASE.
SQ SEQUENCE 3033 AA; 329981 MW; 6B133FED090872B4 CRC64;
```

## Query Match

Best Local Similarity 27.0%; Pred. No. 78;  
Matches 40; Conservative 21; Mismatches 60; Indels 27; Gaps 7;

```
OY 11 VIERYLEKAEKENTTGCAE---HCSLSENIIV-PDTKVFYAKRM-----EV 56
      || : | | : | | | | : | | | | : | | | | : | | | |
Db 1665 VLACGVLAAVAAYCLATGCSITIGRIHLNDQVAVAPDKELVEAFDEMECASKALIBE 1724
OY 57 GQQAWEVWQGLALLSEAVLRGQALLVNSSQWPPEIQ-----LHYDKAVSGLSRLTLLR 110
      || : | | : | | | | : | | | | : | | | | : | | | |
Db 1725 GQMAEMLSKLTGLLQVTKQADIQPMQSSWPXIEQFMARHMMNFTSGIQYLAGIST 1784
OY 111 ALG---AQKEAISPPDAASAPLRTIT 134
Db 1785 LPGNPAVASMAFS---AALTSPLEPTST 1809
```

## RESULT 12

```
Q9FK91 PRELIMINARY; PRT; 815 AA.
ID Q9FK91
AC Q9FK91;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Seed maturation protein Pm38 protein.
DE Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucots II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:203-216(1998).
DR EMBL: AB012244; BAB09119.1; -
DR InterPro: IPR002086; Aldehyde_dehydr.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS50172; BRCT; 1.
SQ SEQUENCE 815 AA; 91534 MW; 4B1B602057D46B60 CRC64;
```

## Query Match

Best Local Similarity 9.7%; Score 82; DB 10; Length 815;  
Matches 42; Conservative 29; Mismatches 52; Indels 70; Gaps 12;

```
OY 21 EAENITTCGA-----EHCISLSENIIVPTKVFYAKRMVEG-----QQAWEVW 64
      | : | | | : | | | | : | | | | : | | | | : | | | |
Db 328 EKDGLLYNCAFSCIDLGKGRNEYC-IMQLVTPDSNLNMY-FKRGVGDPPNAEEELDEM 385
      : : | | : | | : | | | | : | | | | : | | | | : | | | |
OY 65 QGLALLSEAVLRGQALLV-----NSSQWPE-----PLQ-----LHYDAVSG 102
      : : | | : | | : | | | | : | | | | : | | | | : | | | |
Db 386 ED-----EEAAIKERFALFEIAGNEFEPEWRERKKIQKKPKKFPIDMDGIEVRSGLGL 441
```

QY 103 RSL-----TTTLRALGAO-----KEAISPDAASAAPLRTTADTFRL 141  
| : : : : :  
DB 442 RQGIASAHCKLDSEFVANFIKVLCOGEITNYALMELGLDPPD-----LPGMLTDHILKRC 497  
| : : : : :  
QY 142 FRYYSNPNLGGKTK 154  
| : : : : :  
DB 498 EEVLLEFVE-KVK 509  
| : : : : :  
RESULT 13  
O86808  
ID O86808 PRELIMINARY: PRT: 1829 AA.  
AC O86808:  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Putative sensory histidine kinase.  
GN SC05748 OR SC7C7.03.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Harris D., Taylor K.;  
RL Submitted (Jul-1998) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (Jul-1998) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA MEDLINE-97000351, PubMed-8843436;  
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,  
RA Khasn H., Hopwood D.A.;  
RT A set of ordered cosmids and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2) / Chater K.F., Cerdano-Tarraga A.M., Challis G.L.,  
RA Bentley S.D., James K.D., Harris D.E., Quail M.A., Kieser H.,  
RA Thomson N.R., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble L., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Squares R., Taylor K.,  
RA Warren T., Wietorrek A., Woodward J., Barrell B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete Streptomyces  
RT coelicolor A3(2) ".  
RL Nature 417:141-147(2002).  
CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE  
CC KINASES.  
DR EMBL: AL001031; CA19849.1; -.  
DR HSP; P52934; ID23.  
DR InterPro: IPR003594; ATPbind\_ATPase.  
DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
DR InterPro: IPR003018; GAF.  
DR InterPro: IPR003600; HAMF.  
DR InterPro: IPR003661; His\_kinA.  
DR InterPro: IPR004359; His\_kinA.  
DR InterPro: IPR001789; Response\_reg.  
DR Pfam: PF01590; GAF; 1.  
DR Pfam: PF00672; HAMF; 1.  
DR Pfam: PF02518; HAMFase\_C; 1.  
DR Pfam: PF00072; response\_reg; 1.  
DR Pfam: PF00512; signal; 1.  
DR PRINTS: PR00344; BCTRSENSOR.

DR ProDom: PD000039; Response\_reg; 1.  
DR SMART; SM00065; GAF; 1.  
DR SMART; SM00304; HAMF; 12.  
DR SMART; SM00387; HAMFase\_C; 1.  
DR SMART; SM00388; HSKA; 1.  
DR SMART; SM00448; REC; 1.  
DR KINASE: Phosphorylation; Sensory transduction; Transferease.  
SQ SEQUENCE 1829 AA; 195751 MW; C9C869938C956A4 CRC64;  
Query Match 9.4%; Score 79.5; DB 16; Length 1829;  
Best Local Similarity 22.8%; Pred. No. 82;  
Matches 37; Conservative 29; Mismatches 61; Indels 35; Gaps 6;  
QY 18 EAKEENITTCGAHCISINENTVPTKVFAMKRMVEGQAVEWQGLALL----- 70  
| : : : : :  
DB 328 QVREISVTTAVA-NGDLSKRVTP-----ARGEVQLAETINQMTELTIFADEV 377  
| : : : : :  
QY 71 -----SEAVLRGOALLVNSQWPPEPLQLHVDKAVSGRLSTTLRALGAOKEAISP 122  
| : : : : :  
DB 378 TVANETGEGQLGQAVPGAGIWKDL---TDSVNVFNNLTQYKDIAVTTAVASG 434  
| : : : : :  
QY 123 DAASAAPLRTTADTFRLFRYS--NFLRGKALKYGEACR 162  
| : : : : :  
DB 435 DIS-----QKVTVDVAGMELKNTVNTVDLSAFGAEVTR 471  
| : : : : :  
RESULT 14  
Q91ZAI  
ID Q91ZAI PRELIMINARY: PRT: 3033 AA.  
AC Q91ZAI:  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)  
DE (NS1)].  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MD2B-1;  
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,  
RA Sato C.;  
RT "Full length cDNA sequence of HCV genotype 2b, strain MD2B-1.";  
RT Submitted (Feb-2000) to the EMBL/Genbank/DBJ databases.  
CC -!- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
DR EMBL: AF388486; AA59945.1; -.  
DR HSP; P27958; IHE1.  
DR MEROPS; S29.001; -.  
DR InterPro: IPR000345; CytC\_heme\_bind.  
DR InterPro: IPR001410; DEAD.  
DR InterPro: IPR002522; HCV\_capsid.  
DR InterPro: IPR002519; HCV\_core.  
DR InterPro: IPR002531; HCV\_NS1.  
DR InterPro: IPR002518; HCV\_NS2.  
DR InterPro: IPR004109; HCV\_NS3.  
DR InterPro: IPR000745; HCV\_NS4a.  
DR InterPro: IPR001490; HCV\_NS4b.  
DR InterPro: IPR002868; HCV\_NS5a.  
DR InterPro: IPR002166; HCV\_RdRP.  
DR Pfam: PF01543; HCV\_capsid; 1.  
DR Pfam: PF01542; HCV\_core; 1.  
DR Pfam: PF01539; HCV\_env; 1.  
DR Pfam: PF01560; HCV\_NS1; 1.  
DR Pfam: PF01568; HCV\_NS2; 1.  
DR Pfam: PF02907; HCV\_NS3; 1.  
DR Pfam: PF01006; HCV\_NS4a; 1.  
DR Pfam: PF01001; HCV\_NS4b; 1.  
DR Pfam: PF01506; HCV\_NS5a; 1.  
DR Pfam: PF00998; HCV\_RdRP; 1.



DR Prodom: PD186062; HCV NS1. 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KM Polyprotein; RNA-directed RNA polymerase; Transmembrane.  
SQ SEQUENCE 3033 AA; 330723 MW; FE04FEC7C385A13A CRC64;

Query Match 9.4%; Score 79.5; DB 12; Length 3033;  
Best Local Similarity 26.6%; Pred. No. 1.6e+02;  
Matches 41; Conservative 22; Mismatches 64; Indels 27; Gaps 7;

OY 5 LICDSRVLEKYLEAKENITTCAC--HCSLNETIV--PDTKVFYAMKRM----- 54  
DB 1659 IMTSSWVLAGVLAAYCATGICISITGRHLNDQYVAPDKKILYEAEPDEMECASK 1718  
OY 55 ----EVGQAAVEWOG--LALISEAVLRGALLVNSSQPEPLQ---LHYDKAVSGLRS 104  
DB 1719 AALIEGGRMAEMLSKTIQGLQCATRQADIRPAIQSSWPKLEQFMKHMNFISGIQY 1778  
OY 105 LTTLLRAG---AKKEAISPPDASAPLRIT 134  
DB 1779 LAGLSTLPGNPAVVASMAFS---AALISPLPTST 1809

## RESULT 15

O43380 PRELIMINARY; PRT: 480 AA.

AC O43380.  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Putative pp70 ribosomal protein S6 kinase.  
GN ASPK11.  
OS Avena sativa (Oat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Avenae; Avena.  
OX NCBI\_TaxID=4498;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. RHIANNON; TISSUE=ALEURONE;  
RX MEDLINE=95284341; PubMed=7766874;  
RA Huttly A.K., Phillips A.L.;  
RT "qJbberellin regulated expression in oat aleurone cells of two kinases  
RL Plant Mol. Biol. 27:1043-1052(1995).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; X79992; CAA56313.1; -.  
DR HSSP; P05132; ICP.  
DR InterPro: IPR000719; Euk\_Pkinase.  
DR InterPro: IPR000961; Pkinase\_C.  
DR InterPro: IPR002290; Ser\_thr\_Pkinase.  
DR Pfam; PF00069; Pkinase; 1.  
DR Pfam; PF00433; Pkinase\_C; 1.  
DR Prodom; PD000001; Euk\_Pkinase; 1.  
DR SMART; SM00133; S\_TK\_X; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Kinase; Ribosomal protein;  
KW Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 480 AA; 53532 MW; 33596A10IDBB077D CRC64;

Query Match 9.3%; Score 79; DB 10; Length 480;  
Best Local Similarity 22.4%; Pred. No. 17;

Matches 32; Conservative 26; Mismatches 65; Indels 20; Gaps 4;

OY 30 AEHCSLNENITVPDTKVFYAMKREVGQAAVEYWGIALISEA--VLRGALLVNSSOP 87  
DB 51 AVHTPEASTIVTPDSPA-----PLAEGEEYVNDPVITKRSHSLVGPTLLVSQSILP 102  
OY 88 WEPLQLHYDKAVSGLRSLTTLRLAGAKKEAISPPDASAPLRITADTFKRL----- 141

DB 103 LSKLLHSESSSDJ--LECLSKERQSNQALSDDELSTFKNEAVGIDNFEVLYLVGOG 160  
OY 142 --FRVYSNFLRGKILKYTGACR 162  
DB 161 AFGKYYQVRMKGTSELYAMKVMR 183

Search completed: January 2, 2003, 15:12:54  
Job time : 31.4199 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:06:49 ; Search time 32.0967 Seconds  
(without alignments)  
689.156 Million cell updates/sec

Title: US-09-853-731-2

Perfect score: 851

Sequence: 1 APPRLIDSRVLENYLLEAK.....NFLRGKILKLTGACRPTGDR 166

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.101002:\*

1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID       | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 851   | 100.0       | 166    | 8 AAP70398  | Sequence of human  |
| 2          | 851   | 100.0       | 166    | 13 AAR23593 | Recombinant hemato |
| 3          | 851   | 100.0       | 166    | 19 AAW77780 | Human Epo receptor |
| 4          | 851   | 100.0       | 166    | 19 AAW58404 | Human erythropoiet |
| 5          | 851   | 100.0       | 166    | 21 ABB07030 | Modified erythropo |
| 6          | 851   | 100.0       | 166    | 22 AAE02641 | Human erythropoiet |
| 7          | 851   | 100.0       | 166    | 22 AAB66698 | Human erythropoiet |
| 8          | 851   | 100.0       | 166    | 23 ABB77897 | Amino acid sequenc |
| 9          | 851   | 100.0       | 166    | 23 AAM53062 | Human erythropoiet |
| 10         | 851   | 100.0       | 167    | 6 AAP50298  | Human recombinant  |

|    |     |       |     |             |                    |
|----|-----|-------|-----|-------------|--------------------|
| 11 | 851 | 100.0 | 167 | 6 AAP50299  | Human recombinant  |
| 12 | 851 | 100.0 | 169 | 23 ABB77899 | Amino acid sequenc |
| 13 | 851 | 100.0 | 174 | 23 ABB77898 | Amino acid sequenc |
| 14 | 851 | 100.0 | 174 | 23 ABB77900 | Amino acid sequenc |
| 15 | 851 | 100.0 | 188 | 7 AAP60599  | Clone lambda HEPOR |
| 16 | 851 | 100.0 | 188 | 9 AAP81195  | Erythropoietin enc |
| 17 | 851 | 100.0 | 193 | 6 AAP50300  | Human erythropoiet |
| 18 | 851 | 100.0 | 193 | 7 AAP60597  | Clone lambda HEPOR |
| 19 | 851 | 100.0 | 193 | 8 AAP70256  | Sequence of human  |
| 20 | 851 | 100.0 | 193 | 15 AAR65499 | Human prepro-eryth |
| 21 | 851 | 100.0 | 193 | 16 AAR81982 | Human erythropoiet |
| 22 | 851 | 100.0 | 193 | 16 AAR71137 | Human erythropoiet |
| 23 | 851 | 100.0 | 193 | 16 AAR74141 | Human erythropoiet |
| 24 | 851 | 100.0 | 193 | 17 AAR98397 | Human erythropoiet |
| 25 | 851 | 100.0 | 193 | 21 AAY94530 | Human erythropoiet |
| 26 | 851 | 100.0 | 193 | 21 AAY93638 | Amino acid sequenc |
| 27 | 851 | 100.0 | 193 | 21 AAY99704 | Human non-glycosyl |
| 28 | 851 | 100.0 | 193 | 21 AAY43398 | Human erythropoiet |
| 29 | 851 | 100.0 | 193 | 22 AAB85573 | Human erythropoiet |
| 30 | 851 | 100.0 | 193 | 22 AAB34978 | Human erythropoiet |
| 31 | 851 | 100.0 | 193 | 23 AAE15341 | Human erythropoiet |
| 32 | 851 | 100.0 | 194 | 16 AAR71167 | Human erythropoiet |
| 33 | 851 | 100.0 | 194 | 19 AAW62048 | Human erythropoiet |
| 34 | 851 | 100.0 | 194 | 21 AAB10654 | Human erythropoiet |
| 35 | 851 | 100.0 | 196 | 23 ABB77902 | Amino acid sequenc |
| 36 | 851 | 100.0 | 201 | 23 ABB77901 | Amino acid sequenc |
| 37 | 851 | 100.0 | 201 | 23 ABB77903 | Amino acid sequenc |
| 38 | 851 | 100.0 | 302 | 13 AAR23596 | Recombinant hemato |
| 39 | 851 | 100.0 | 303 | 13 AAR23598 | Recombinant hemato |
| 40 | 851 | 100.0 | 321 | 13 AAR23075 | IL-3:Epo short, re |
| 41 | 851 | 100.0 | 321 | 13 AAR23597 | Recombinant hemato |
| 42 | 851 | 100.0 | 322 | 13 AAR23599 | Epo:IL-3 short, re |
| 43 | 851 | 100.0 | 330 | 13 AAR23076 | Epo:IL-3 short, re |
| 44 | 851 | 100.0 | 340 | 13 AAR23078 | IL-3:Epo Flex, rec |
| 45 | 851 | 100.0 | 349 | 13 AAR23079 | Epo:IL-3 Flex, rec |

#### ALIGNMENTS

RESULT 1

AAAP70398 standard; protein; 166 AA.

AC AAP70398;

XX

XX

DT 19-FEB-1991 (first entry)

XX

DE Sequence of human erythropoietin (Epo).

XX

XX Mega-karyocyte-platelet growth factor; hormone;

KW mega-karyocyte colony stimulating factor; therapy;

KW small acetyl cholinesterase positive cell;

KW erythrocyte growth effect.

XX

OS Homo sapiens.

XX

PN JP62149624-A.

XX

PD 03-JUL-1987.

XX

PF 15-AUG-1986; 86JP-0191542.

XX

PR 13-SEP-1985; 85JP-0203049.

XX

PR 15-AUG-1986; 86JP-0191542.

XX

PA (KAWA/) KAWAKITA M.

XX

DR WPI; 1987-224837/32.

XX

PT Mega-karyocyte-platelet growth factor - contains as active

PT component human erythropoietin and is used to treat diseases

PT caused by decrease in platelets

XX Disclosure; Page 181; 8pp; Japanese.  
 PS  
 CC All of the Cys residues in the SQ are labelled "SH". Megakaryocyte-  
 CC Platelet growth factor contains human EPO as an active principle.  
 CC Human EPO has a megakaryocyte colony-stimulating activity and  
 CC increases the ratio of small acetyl cholinesterase positive and  
 CC (S&ChE-) which is immature megakaryocyte. Human EPO effects  
 CC megakaryocyte-platelet system other than an erythrocyte growth  
 CC effect. Megakaryocyte-platelet growth is usable as a remedy for  
 CC diseases caused by a platelet decrease.  
 CC  
 SQ Sequence 166 AA;

Query Match 100.0%; Score 851; DB 8; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-87;  
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 APPRLICDSRYLERLYLEAKFAENITTCGAHCGLNENITVPDKVNFYAKRMVEGQQA 60  
 DB 1 APPRLICDSRYLERLYLEAKFAENITTCGAHCGLNENITVPDKVNFYAKRMVEGQQA 60  
 QY 61 VEWOGIALLSEAVLRGQALLVNSQPEPQLHVDKAVSGLSRLTTLRLAGAKRAIS 120  
 DB 61 VEWOGIALLSEAVLRGQALLVNSQPEPQLHVDKAVSGLSRLTTLRLAGAKRAIS 120  
 QY 121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLYTGECRCRCDR 166  
 DB 121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLYTGECRCRCDR 166

RESULT 2  
 AAR23593  
 ID AAR23593 standard; Protein; 166 AA.  
 AC AAR23593;  
 XX  
 DT 20-OCT-1992 (first entry)  
 XX  
 DE Recombinant hematopoietic molecule portion 2.  
 XX  
 KW Erythropoietin; EPO; erythrocytes; IL-3; haematopoiesis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9206116-A.  
 XX  
 PD 16-APR-1992.  
 XX  
 PF 26-SEP-1991; 91WO-US07053.  
 XX  
 PR 28-SEP-1990; 90US-058958.  
 XX  
 PA (ORTH ) ORTHO PHARM CORP.  
 XX  
 PI Rosen JI;  
 XX  
 DR WPI; 1992-150819/18.  
 XX  
 PT Recombinant haematopoietic molecules useful in treating  
 PT anemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-3 or M-CSF  
 PT and has early and later myeloid differentiation activity  
 XX  
 PS Disclosure; Page 32; 82pp; English.  
 XX

CC This protein sequence given comprises the entire amino acid sequence  
 CC of human erythropoietin (EPO). EPO leads to the maturation of  
 CC erythrocytes and is therefore designated as a late myeloid  
 CC differentiation factor (MDF). Within the scope of the invention  
 CC hybrid molecules were produced which contain at least a portion of an  
 CC early MDF and at least a portion of a late MDF covalently linked. The  
 CC EPO sequence given is effective within the scope of the invention in  
 CC full or in a truncated version. Amino acids 7-161 act as a

CC Late MDF when recombined with an early MDF eg. IL-3.  
 CC These compounds can be used to promote hematopoiesis in a patient.  
 CC The bonding of the early and late factors allows a very high conc. of  
 CC late MDF at the surface of a cell which the early MDF is bound. It  
 CC also allows the early MDF to act more specifically to stimulate only  
 CC the desired lineage, thus reducing undesirable effects. These  
 CC compounds are useful for treating anemias of various origins eg. renal  
 CC failure and AIDS. It is easier to produce and administer one  
 CC recombinant molecule rather than two separate molecules.  
 CC  
 SQ Sequence 166 AA;

Query Match 100.0%; Score 851; DB 13; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-87;  
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 APPRLICDSRYLERLYLEAKFAENITTCGAHCGLNENITVPDKVNFYAKRMVEGQQA 60  
 DB 1 APPRLICDSRYLERLYLEAKFAENITTCGAHCGLNENITVPDKVNFYAKRMVEGQQA 60  
 QY 61 VEWOGIALLSEAVLRGQALLVNSQPEPQLHVDKAVSGLSRLTTLRLAGAKRAIS 120  
 DB 61 VEWOGIALLSEAVLRGQALLVNSQPEPQLHVDKAVSGLSRLTTLRLAGAKRAIS 120  
 QY 121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLYTGECRCRCDR 166  
 DB 121 PPDASAAPLRTITADTFPRKLFYRYSNFLRGKIKLYTGECRCRCDR 166

RESULT 3  
 AAW77780  
 ID AAW77780 standard; Protein; 166 AA.  
 AC AAW77780;  
 XX  
 DT 24-NOV-1998 (first entry)  
 XX  
 DE Human EPO receptor agonist polypeptide.  
 XX  
 KW Haematopoietic receptor agonist; erythropoietin receptor agonist;  
 KW EPO; human; chimeric protein; stem cell expansion; tumour;  
 KW infection; autoimmune disease; haematopoietic disorder; therapy;  
 KW dendritic cell.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..6 /note= "1-6 amino acids of the N-terminus are  
 FT Misc-difference 162..166 optionally deleted"  
 FT Misc-difference 162..166 /note= "1-5 amino acids of the C-terminus are  
 FT Misc-difference 23..24 optionally deleted"  
 FT Misc-difference 23..24 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 24..25 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 25..26 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 26..27 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 27..28 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 28..29 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 29..30 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 30..31 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 31..32 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 32..33 /note= "possible positions of new C- and N-termini"  
 FT Misc-difference 32..33 /note= "possible positions of new C- and N-termini"



CC R1-L1-R2, R2-L1-R1, R1-R2 or R2-R1, where L is a linker and R1 and  
 CC R2 are independently selected from: (a) the human EPO receptor  
 CC agonist; (b) a human stem cell factor receptor agonist polypeptide  
 CC (see AAW77781); (c) a human flt-3 receptor agonist polypeptide (see  
 CC AAW77782); (d) a modified human granulocyte colony stimulating factor  
 CC (G-CSF) polypeptide (see AAW77783); (e) modified human interleukin-3  
 CC polypeptide (see AAW77784); (f) modified human c-mpl ligand polypeptide  
 CC (see AAW77785); and (g) a factor selected from the group consisting of  
 CC a CSF, a cytokine, a lymphokine, an interleukin and a haematopoietic  
 CC growth factor, provided that at least R1 or R2 is selected from (a),  
 CC (b) or (c) as above. The multi-functional chimeric haematopoietic  
 CC receptor agonist can be used to stimulate the production of

Query Match 100.0%; Score 851; DB 19; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-87;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRYLERYLEAKAEENITTCGAHCSSINENITVPDTKYNFYAMKMEYGOOA 60  
 Db 1 APPRLICDSRYLERYLEAKAEENITTCGAHCSSINENITVPDTKYNFYAMKMEYGOOA 60  
 QY 61 VEWOGIALSEAVLRGOALLVNSOPWEPLQIHVDKAVSGLSLTLTLRALCAQKEAIS 120  
 Db 61 VEWOGIALSEAVLRGOALLVNSOPWEPLQIHVDKAVSGLSLTLTLRALCAQKEAIS 120  
 QY 121 PPDAAASAPLRTITADTFKRLFRVYSNFIKLGKLTGTGACRTGDR 166  
 Db 121 PPDAAASAPLRTITADTFKRLFRVYSNFIKLGKLTGTGACRTGDR 166

## RESULT 4

AAW58404  
 ID AAW58404 standard; Protein; 166 AA.

AC AAW58404;

DT 12-OCT-1998 (first entry)

DE Human erythropoietin.

KW Erythropoietin receptor agonist; EPO; human; anaemia;

KW haematopoietic deficiency; red blood cell; erythroid progenitor;

KW bone marrow suppression.

OS Homo sapiens.

PN WO9818926-A1.

PD 07-MAY-1998.

PE 23-OCT-1997; 97WO-US18703.

PR 25-OCT-1996; 96US-0034044.

PA (SEAR ) SEARLE & CO G D.

PI Feng Y, McWharter CA, Summers N;

DR WPI: 1998-272221/24.

PT N-PSDB; AAV31031.

Human erythropoietin receptor agonist polypeptide - used to

stimulate the production of red blood cells in a patient

Claim 1; Page 93; 112pp; English.

CC A claimed human erythropoietin (EPO) receptor agonist polypeptide  
 CC comprises a modified EPO amino acid sequence given in AAW58404,  
 CC where (a) optionally 1-6 amino acids from the N-terminus and 1-5  
 CC from the C-terminus can be deleted, (b) the N-terminus is joined to  
 CC the C-terminus directly or through a linker (see AAW58405-12) capable  
 CC of joining the N-terminus to the C-terminus, (c) there are new C-  
 CC and N-terminal at any two consecutive amino acids from amino acids

CC 23-24 to 38-39, 40-41 to 41-42, 43-44 to 48-49, 50-51 to 57-58,  
 CC 77-78 to 82-83, 84-85 to 88-89, and 108-109 to 131-132, and (d)  
 CC optionally the agonist polypeptide is preceded by Met, Ala, or  
 CC Met-Ala. 60 Of these circularly permuted EPO receptor agonists  
 CC (see AAW58413-72) are claimed. Also claimed are: nucleic acid  
 CC molecules (see AAV30971-V31030) encoding novel EPO receptor agonists;  
 CC a method of producing an EPO receptor agonist using transformed or  
 CC transfected host cells; and methods for stimulating the production  
 CC of haematopoietic cells, for selective ex vivo expansion of  
 CC erythroid progenitors, and treating patients having a haematopoietic  
 CC disorder using the EPO receptor agonists. The EPO receptor  
 CC agonists use one or more activities of native EPO and may also  
 CC show improved haematopoietic cell-stimulating activity and/or an  
 CC improved biological activity profile which may include reduction of undesirable  
 CC biological activities associated with native EPO and/or have  
 CC improved physical properties such as increased solubility,  
 CC stability and refold efficiency.

Sequence 166 AA:

Query Match 100.0%; Score 851; DB 19; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-87;  
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRYLERYLEAKAEENITTCGAHCSSINENITVPDTKYNFYAMKMEYGOOA 60  
 Db 1 APPRLICDSRYLERYLEAKAEENITTCGAHCSSINENITVPDTKYNFYAMKMEYGOOA 60  
 QY 61 VEWOGIALSEAVLRGOALLVNSOPWEPLQIHVDKAVSGLSLTLTLRALCAQKEAIS 120  
 Db 61 VEWOGIALSEAVLRGOALLVNSOPWEPLQIHVDKAVSGLSLTLTLRALCAQKEAIS 120  
 QY 121 PPDAAASAPLRTITADTFKRLFRVYSNFIKLGKLTGTGACRTGDR 166  
 Db 121 PPDAAASAPLRTITADTFKRLFRVYSNFIKLGKLTGTGACRTGDR 166

## RESULT 5

ABB07030  
 ID ABB07030 standard; Protein; 166 AA.

AC ABB07030;

DT 21-JUN-2002 (first entry)

DE Modified erythropoietin related gene protein sequence.

KW Modified erythropoietin; EPO.

OS Unidentified.

PN KR145802-B1.

PD 01-AUG-1998.

PE 31-MAY-1994; 94KR-0012082.

PR 31-MAY-1994; 94KR-0012082.

PA (GLDS ) LG CHEM CO LTD.

PI Kim C, Song Y, Lee T;

DR WPI: 2000-234250/20.

PT N-PSDB; ABL50878.

MODIFIED ERYTHROPOIETIN GENE AND EXPRESSION VECTORS THEREOF -

Disclosure; Page 14; 15pp; Korean.

CC The present invention describes modified erythropoietin (EPO) genes  
 CC and expression vectors comprising the genes. The present sequence  
 CC represents a protein sequence from the present invention.



```

Db 1 APRRLICDSRVLEIRYLLEAKENITTTGCAEHCSINENTIVPDTXVNFYAKRMVEVGOA 60
QY 61 VEWOGIALISEAVLRGQALVNSSQPWEPQLQHYDKANVSGIRSLTTLLRALGAQKEAIS 120
Db 61 VEWOGIALISEAVLRGQALVNSSQPWEPQLQHYDKANVSGIRSLTTLLRALGAQKEAIS 120
QY 121 PPDASAAPLRTITADTFRKLFRRVYSNFLRGKIKLYTGACRTGDR 166
Db 121 PPDASAAPLRTITADTFRKLFRRVYSNFLRGKIKLYTGACRTGDR 166

```

## RESULT 8

ABB77897 ID ABB77897 standard; protein: 166 AA.

AC ABB77897;

DT 07-OCT-2002 (first entry)

XX Amino acid sequence of a human erythropoietin (EPO).

XX Human; erythropoietin; EPO; glycoprotein; reticulocyte production;

KW red blood cell production; anaemia; chronic renal failure;

KM acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;

XX committed erythroid progenitor.

XX Homo sapiens.

PN WO200249673-A2.

XX 27-JUN-2002.

XX 08-DEC-2001; 2001WO-EP14434.

XX 20-DEC-2000; 2000EP-0127891.

XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.

PI Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischer W;

PI Wozny M;

DR WPI: 2002-566640/60.

XX Novel conjugate of erythropoietin glycoprotein with polyethylene

PT glycol, useful for treating diseases correlated with anaemia in chronic

PT renal failure patients and acquired immunodeficiency syndrome -

XX Claim 26; Fig 2; 40pp; English.

XX The present sequence represents a human erythropoietin (EPO) protein.

XX It was used to produce conjugates of the invention. The specification

XX describes a conjugate comprising an EPO glycoprotein having an N-terminal

XX alpha-amino group, chosen from human EPO (hEPO) or its analogues (where

XX hEPO is modified by addition of 1-6 glycosylation sites or a

XX rearrangement of a glycosylation site). The glycoprotein is covalently

XX linked to a poly(ethylene glycol) group. The EPO glycoprotein has in vivo

XX biological activity of causing bone marrow cells to increase production

XX of reticulocytes and red blood cells. The conjugate increased circulating

XX half-life and plasma residence time, decreased clearance, increased

XX clinical activity in vivo, improved potency and stability, when compared

XX for the treatment and prophylaxis of diseases correlated with anaemia in

XX chronic renal failure patients (CRF), acquired immunodeficiency syndrome

XX (AIDS) and for treating cancer patients undergoing chemotherapy. It is

XX also useful for treating patients by stimulating the division and

XX differentiation of committed erythroid progenitors in the bone marrow.

XX Sequence 166 AA;

Query Match 100.0%; Score 851; DB 23; Length 166;

Best Local Similarity 100.0%; Pred. No. 4, 3e-87;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 APRRLICDSRVLEIRYLLEAKENITTTGCAEHCSINENTIVPDTXVNFYAKRMVEVGOA 60
Db 1 APRRLICDSRVLEIRYLLEAKENITTTGCAEHCSINENTIVPDTXVNFYAKRMVEVGOA 60
QY 61 VEWOGIALISEAVLRGQALVNSSQPWEPQLQHYDKANVSGIRSLTTLLRALGAQKEAIS 120
Db 61 VEWOGIALISEAVLRGQALVNSSQPWEPQLQHYDKANVSGIRSLTTLLRALGAQKEAIS 120
QY 121 PPDASAAPLRTITADTFRKLFRRVYSNFLRGKIKLYTGACRTGDR 166
Db 121 PPDASAAPLRTITADTFRKLFRRVYSNFLRGKIKLYTGACRTGDR 166

```

## RESULT 9

AAM53062 ID AAM53062 standard; protein: 166 AA.

XX AAM53062;

DT 25-MAR-2002 (first entry)

XX Human erythropoietin (hEPO), 166 residue form.

XX Human; erythropoietin; EPO; hEPO; haemostatic; red blood cell;

KW blood disorder; anaemia; chronic renal failure; CRF; AIDS;

KM acquired immunodeficiency syndrome; cancer chemotherapy; haemostatic;

XX anti-RTV; antianaemic.

XX Homo sapiens.

OS Key Location/Qualifiers

XX Disulfide-bond 7..161

XX Modified-site 24

XX /note= "N-glycosylated"

XX Disulfide-bond 29..33

XX Modified-site 38

XX /note= "N-glycosylated"

XX Modified-site 83

XX /note= "N-glycosylated"

XX Modified-site 126

XX /note= "O-glycosylated"

XX WO200187329-A1.

XX 22-NOV-2001.

XX 08-MAY-2001; 2001WO-EP05187.

XX 15-MAY-2000; 2000EP-0110355.

XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.

XX Papadimitriou A;

XX WPI: 2002-082943/11.

XX Composition useful in the treatment of e.g. AIDS comprises an

XX erythropoietin protein, and a multiple charged inorganic anion in a

XX buffer -

XX Claim 28; Fig 2; 64pp; English.

XX The invention relates to liquid pharmaceutical compositions comprising

XX an erythropoietin (EPO) protein, a multiple negatively charged inorganic

XX anion in a buffer which maintains the pH of the solution from 5.5-7.0,

XX and optionally at least one excipient. The erythropoietin used in the

XX composition is preferably human (AAM53061 or AAM53062) a human

XX (AAM53061-AAM53107), or an erythropoietin with the C-terminal addition

XX of a C-terminal fragment of human chorionic gonadotropin (AAM53063).

XX Erythropoietin is a glycoprotein essential for the formation of red blood

XX cells and is therefore useful in the treatment of blood disorders



CC characterised by low or defective red blood cell production. The  
CC compositions of the invention can be used in the treatment and prevention  
CC of anaemia in chronic renal failure patients (CRF), AIDS (acquired  
CC immunodeficiency syndrome), and/or for the treatment of cancer patients  
CC undergoing chemotherapy. Unlike prior art erythropoietin compositions,  
CC the compositions of the invention do not contain human serum albumin  
CC (thereby avoiding the possibility of viral infections and allergic  
CC reactions associated with this component), are liquid rather than  
CC lyophilisates (and therefore do not need to be reconstituted before  
CC administration), and are stable at elevated temperatures such as 25  
CC degrees Celsius and even 40 degrees Celsius, and therefore can be stored  
CC without refrigeration for prolonged periods without degradation and loss  
CC of activity. The present sequence represents the 166 residue form  
CC of human erythropoietin which is specifically claimed for use in a  
CC composition of the invention.  
CC  
XX Sequence 166 AA;  
SQ  
Query Match 100.0%; Score 851; DB 23; Length 166;  
Best Local Similarity 100.0%; Pred. No. 4.3e-87;  
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 APPRLICDSRVLEKRLLEAKAEENITTCGAHCNSINENITVPDTRKVNFRYAKRMVEVGQA 60  
Db 1 APPRLICDSRVLEKRLLEAKAEENITTCGAHCNSINENITVPDTRKVNFRYAKRMVEVGQA 60  
Qy 61 VEWMOGLALISEAVIRGQALLVNSSQPEWPIQLHVDKAVSGRLSTTLRALGAQKEAIS 120  
Db 61 VEWMOGLALISEAVIRGQALLVNSSQPEWPIQLHVDKAVSGRLSTTLRALGAQKEAIS 120  
Qy 121 PPDAASAAPLRTITADTRFKLFRRVYSNPLRGKRLKLYTGECRTGDR 166  
Db 121 PPDAASAAPLRTITADTRFKLFRRVYSNPLRGKRLKLYTGECRTGDR 166  
RESULT 10  
AAP50298  
ID AAP50298 standard; protein: 167 AA.  
XX  
AC AAP50298;  
XX  
DT 01-JAN-1980 (first entry)  
XX  
DE Human recombinant erythropoietin expressed in *Saccharomyces*  
DE *cerevisiae*.  
XX  
KW Erythropoietin; red blood cell; erythrocyte; anaemia; blood;  
KW disorder; ds; *Saccharomyces cerevisiae*.  
XX  
OS Homo sapiens.  
XX  
PN W08502610-A.  
XX  
PD 20-JUN-1985.  
XX  
PE 11-DEC-1984; 84WO-US02021.  
XX  
PR 30-NOV-1984; 84US-0675298.  
PR 13-DEC-1983; 83US-0561024.  
PR 21-FEB-1984; 84US-0582185.  
PR 28-SEP-1984; 84US-0655841.  
XX  
PA (KIRI-) KIRIN-AMGEN INC.  
XX  
XX WPI: 1985-159229/26.  
DR N-PSDB: AAN50345.  
XX  
XX New polypeptide having properties of erythropoietin - is prepd.  
PT by cultivation of transformed eucaryotic or procaryotic host  
PS Disclosure; Page 82; 113pp; English.  
XX  
XX Human erythropoietin encoded by this sequence is essential for red

CC blood cell formation and is used for the diagnosis and treatment of  
CC blood disorders such as anaemia. Large amounts of EPO may be obtained  
CC using recombinant DNA techniques in contrast to small amounts  
CC obtained from plasma and urine. This sequence is expressed in *S.*  
CC *cerevisiae*. See also AAN50346-50 and AAP50299-P50301.  
XX  
XX Sequence 167 AA;  
SQ  
Query Match 100.0%; Score 851; DB 6; Length 167;  
Best Local Similarity 100.0%; Pred. No. 4.3e-87;  
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 APPRLICDSRVLEKRLLEAKAEENITTCGAHCNSINENITVPDTRKVNFRYAKRMVEVGQA 60  
Db 2 APPRLICDSRVLEKRLLEAKAEENITTCGAHCNSINENITVPDTRKVNFRYAKRMVEVGQA 61  
Qy 61 VEWMOGLALISEAVIRGQALLVNSSQPEWPIQLHVDKAVSGRLSTTLRALGAQKEAIS 120  
Db 62 VEWMOGLALISEAVIRGQALLVNSSQPEWPIQLHVDKAVSGRLSTTLRALGAQKEAIS 121  
Qy 121 PPDAASAAPLRTITADTRFKLFRRVYSNPLRGKRLKLYTGECRTGDR 166  
Db 122 PPDAASAAPLRTITADTRFKLFRRVYSNPLRGKRLKLYTGECRTGDR 167  
RESULT 11  
AAP50299  
ID AAP50299 standard; protein: 167 AA.  
XX  
AC AAP50299;  
XX  
DT 01-JAN-1980 (first entry)  
XX  
DE Human recombinant erythropoietin expressed in *Escherichia coli*.  
DE  
XX  
KW Erythropoietin; red blood cell; erythrocyte; anaemia; blood;  
KW disorder; ds; *Escherichia coli*.  
XX  
OS Homo sapiens.  
XX  
PN W08502610-A.  
XX  
PD 20-JUN-1985.  
XX  
PE 11-DEC-1984; 84WO-US02021.  
XX  
PR 30-NOV-1984; 84US-0675298.  
PR 13-DEC-1983; 83US-0561024.  
PR 21-FEB-1984; 84US-0582185.  
PR 28-SEP-1984; 84US-0655841.  
XX  
PA (KIRI-) KIRIN-AMGEN INC.  
XX  
XX WPI: 1985-159229/26.  
DR N-PSDB: AAN50346.  
XX  
XX New polypeptide having properties of erythropoietin - is prepd.  
PT by cultivation of transformed eucaryotic or procaryotic host  
PS Disclosure; Page 72; 113pp; English.  
XX  
XX Human erythropoietin encoded by this sequence is essential for red  
CC blood cell formation and is used for the diagnosis and treatment of  
CC blood disorders such as anaemia. Large amounts of EPO may be obtained  
CC using recombinant DNA techniques in contrast to small amounts  
CC obtained from plasma and urine. This sequence is expressed in *E.*  
CC *coli*. See also AAN50345, AAN50347-50 and AAP50298, AAP50300-P50301.  
XX  
XX Sequence 167 AA;  
SQ  
Query Match 100.0%; Score 851; DB 6; Length 167;  
Best Local Similarity 100.0%; Pred. No. 4.3e-87;  
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRYLERLYLEAKEAENITTCGAHCSLNENITVPDTRKVNPFYAKRMREVGQA 60  
 |||  
 DB 2 APPRLICDSRYLERLYLEAKEAENITTCGAHCSLNENITVPDTRKVNPFYAKRMREVGQA 61  
 |||  
 QY 61 VEWMOGLALISEAVLRGALLVNSSQPEPIQLHVDKAVSGLRSITLLIRALGAQKEAIS 120  
 |||  
 DB 62 VEWMOGLALISEAVLRGALLVNSSQPEPIQLHVDKAVSGLRSITLLIRALGAQKEAIS 121  
 |||  
 QY 121 PPDASAAPLRTITADTRFKLFRRVSNFLRGKIKLYTGEACRTGDR 166  
 |||  
 DB 122 PPDASAAPLRTITADTRFKLFRRVSNFLRGKIKLYTGEACRTGDR 167  
 |||

## RESULT 12

ID ABB77899 standard; protein; 169 AA.

AC ABB77899;

DT 07-OCT-2002 (first entry)

DE Amino acid sequence of a modified human erythropoietin (EPO).

KW Human: erythropoietin; EPO; glycoprotein; reticulocyte production;

KW red blood cell production; anaemia; chronic renal failure;

KW acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;

XX committed erythroid progenitor.

OS Synthetic.

XX Homo sapiens.

XX Key

FT Cleavage-site 1..3

FT Protein 4..174

FT Protein /note- "EPO protein"

PN MO200249673-A2.

PD 27-JUN-2002.

PE 08-DEC-2001; 2001WO-EP14434.

PR 20-DEC-2000; 2000EP-0127891.

PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.

PI Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischer W;

PI Wozny M;

XX WPI; 2002-566640/60.

XX Novel conjugate of erythropoietin glycoprotein with polyethylene

PT glycol, useful for treating diseases correlated with anaemia in chronic

PT renal failure patients and acquired immunodeficiency syndrome -

XX Disclosure; Page 39; 40pp; English.

XX The present sequence represents a modified human erythropoietin (EPO)

XX protein. The EPO was extended at the N-terminal by a proteolytic

XX cleavage site. It was used to produce conjugates of the invention. The

XX specification describes a conjugate comprising an EPO glycoprotein having

XX an N-terminal alpha-amino group, chosen from human EPO (hEPO) or its

XX analogues (where hEPO is modified by addition of 1-6 glycosylation sites

XX or a rearrangement of a glycosylation site). The glycoprotein is

XX covalently linked to a poly(ethylene glycol) group. The EPO glycoprotein

XX has in vivo biological activity of causing bone marrow cells to increase

XX production of reticulocytes and red blood cells. The conjugate increased

XX circulating half-life and plasma residence time, decreased clearance

XX compared to unmodified EPO. The EPO conjugate is useful for preparing

XX medicaments for the treatment and prophylaxis of diseases correlated with

CC anaemia in chronic renal failure patients (CRF), acquired

CC immunodeficiency syndrome (AIDS) and for treating cancer patients

CC undergoing chemotherapy. It is also useful for treating patients by

CC stimulating the division and differentiation of committed erythroid

CC progenitors in the bone marrow.

XX Sequence 169 AA:

QY 1 APPRLICDSRYLERLYLEAKEAENITTCGAHCSLNENITVPDTRKVNPFYAKRMREVGQA 60  
 |||  
 DB 4 APPRLICDSRYLERLYLEAKEAENITTCGAHCSLNENITVPDTRKVNPFYAKRMREVGQA 63  
 |||  
 DB 61 VEWMOGLALISEAVLRGALLVNSSQPEPIQLHVDKAVSGLRSITLLIRALGAQKEAIS 120  
 |||  
 DB 64 VEWMOGLALISEAVLRGALLVNSSQPEPIQLHVDKAVSGLRSITLLIRALGAQKEAIS 123  
 |||  
 QY 121 PPDASAAPLRTITADTRFKLFRRVSNFLRGKIKLYTGEACRTGDR 166  
 |||  
 DB 124 PPDASAAPLRTITADTRFKLFRRVSNFLRGKIKLYTGEACRTGDR 169  
 |||

## RESULT 13

ID ABB77898 standard; protein; 174 AA.

AC ABB77898;

DT 07-OCT-2002 (first entry)

DE Amino acid sequence of a modified human erythropoietin (EPO).

KW Human: erythropoietin; EPO; glycoprotein; reticulocyte production;

KW red blood cell production; anaemia; chronic renal failure;

KW acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;

XX committed erythroid progenitor.

OS Synthetic.

XX Homo sapiens.

XX Key

FT Cleavage-site 1..8

FT Protein 9..174

FT Protein /note- "EPO protein"

PN MO200249673-A2.

PD 27-JUN-2002.

PE 08-DEC-2001; 2001WO-EP14434.

PR 20-DEC-2000; 2000EP-0127891.

PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.

PI Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischer W;

PI Wozny M;

XX WPI; 2002-566640/60.

XX Novel conjugate of erythropoietin glycoprotein with polyethylene

PT glycol, useful for treating diseases correlated with anaemia in chronic

PT renal failure patients and acquired immunodeficiency syndrome -

XX Disclosure; Page 38-39; 40pp; English.

XX The present sequence represents a modified human erythropoietin (EPO)

XX protein. The EPO was extended at the N-terminal by a proteolytic

XX cleavage site. It was used to produce conjugates of the invention. The

CC specification describes a conjugate comprising an EPO glycoprotein having  
CC an N-terminal alpha-amino group, chosen from human EPO (hEPO) or its  
CC analogues (where hEPO is modified by addition of 1-6 glycosylation sites  
CC or a rearrangement of a glycosylation site). The glycoprotein is  
CC covalently linked to a poly(ethylene glycol) group. The EPO glycoprotein  
CC has *in vivo* biological activity of causing bone marrow cells to increase  
CC production of reticulocytes and red blood cells. The conjugate increased  
CC circulating half-life and plasma residence time, decreased clearance,  
CC increased clinical activity *in vivo*, improved potency and stability, when  
CC compared to unmodified EPO. The EPO conjugate is useful for preparing  
CC medicaments for the treatment and prophylaxis of diseases correlated with  
CC anaemia in chronic renal failure patients (CRF), acquired  
CC immunodeficiency syndrome (AIDS) and for treating cancer patients  
CC undergoing chemotherapy. It is also useful for treating patients by  
CC stimulating the division and differentiation of committed erythroid  
CC progenitors in the bone marrow.

SQ Sequence 174 AA;

Query Match 100.0%; Score 851; DB 23; Length 174;  
Best Local Similarity 100.0%; Pred. No. 4,6e-87;  
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCSSLNENITVPDTRKVFYAMKMEVGOQA 60  
DB 9 APPRLICDSRVLEERYLLEAKAEENITTCGAHCSSLNENITVPDTRKVFYAMKMEVGOQA 68

QY 61 VEVWOGLLLSFAYLRGQALLVNSSQPEPIQLHVDKAVSGLRSLTTLRALGAQKEAIS 120  
DB 69 VEVWOGLLLSFAYLRGQALLVNSSQPEPIQLHVDKAVSGLRSLTTLRALGAQKEAIS 128

QY 121 PPDASAAPLRTITADTFPKLFRVYSNPLRGKILKLTGTGACRTGDR 166  
DB 129 PPDASAAPLRTITADTFPKLFRVYSNPLRGKILKLTGTGACRTGDR 174

RESULT 14  
ABB77900  
ID ABB77900 standard; protein; 174 AA.

AC ABB77900;

DT 07-OCT-2002 (first entry)

DE Amino acid sequence of a modified human erythropoietin (EPO).

KM Human; erythropoietin; EPO; glycoprotein; reticulocyte production;

KW red blood cell production; anaemia; chronic renal failure;

KM acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;

KW committed erythroid progenitor.

OS Synthetic.

OS Homo sapiens.

Key Location/Qualifiers

FT Cleavage-site 1..8 /note= "proteolytic cleavage site"

FT Protein 9..174 /note= "EPO protein"

MO200249673-A2.

27-JUN-2002.

08-DEC-2001; 2001WO-EP14434.

20-DEC-2000; 2000EP-0127891.

(HOFF ) HOFFMANN LA ROCHE & CO AG F.

Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischner W,

Wozny M;

DR WPI; 2002-566640/60.

PT Novel conjugate of erythropoietin glycoprotein with polyethylene  
PT glycol, useful for treating diseases correlated with anaemia in chronic  
PT renal failure patients and acquired immunodeficiency syndrome -

XX Disclosure; Page 39-40; 40pp; English.

CC The present sequence represents a modified human erythropoietin (EPO)  
CC protein. The EPO was extended at the N-terminal by a proteolytic  
CC cleavage site. It was used to produce conjugates of the invention. The  
CC specification describes a conjugate comprising an EPO glycoprotein having  
CC an N-terminal alpha-amino group, chosen from human EPO (hEPO) or its  
CC analogues (where hEPO is modified by addition of 1-6 glycosylation sites  
CC or a rearrangement of a glycosylation site). The glycoprotein is  
CC covalently linked to a poly(ethylene glycol) group. The EPO glycoprotein  
CC has *in vivo* biological activity of causing bone marrow cells to increase  
CC production of reticulocytes and red blood cells. The conjugate increased  
CC circulating half-life and plasma residence time, decreased clearance,  
CC increased clinical activity *in vivo*, improved potency and stability, when  
CC compared to unmodified EPO. The EPO conjugate is useful for preparing  
CC medicaments for the treatment and prophylaxis of diseases correlated with  
CC anaemia in chronic renal failure patients (CRF), acquired  
CC immunodeficiency syndrome (AIDS) and for treating cancer patients  
CC undergoing chemotherapy. It is also useful for treating patients by  
CC stimulating the division and differentiation of committed erythroid  
CC progenitors in the bone marrow.

SQ Sequence 174 AA;

Query Match 100.0%; Score 851; DB 23; Length 174;  
Best Local Similarity 100.0%; Pred. No. 4,6e-87;  
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCSSLNENITVPDTRKVFYAMKMEVGOQA 60  
DB 9 APPRLICDSRVLEERYLLEAKAEENITTCGAHCSSLNENITVPDTRKVFYAMKMEVGOQA 68

QY 61 VEVWOGLLLSFAYLRGQALLVNSSQPEPIQLHVDKAVSGLRSLTTLRALGAQKEAIS 120  
DB 69 VEVWOGLLLSFAYLRGQALLVNSSQPEPIQLHVDKAVSGLRSLTTLRALGAQKEAIS 128

QY 121 PPDASAAPLRTITADTFPKLFRVYSNPLRGKILKLTGTGACRTGDR 166  
DB 129 PPDASAAPLRTITADTFPKLFRVYSNPLRGKILKLTGTGACRTGDR 174

RESULT 15  
AAP60599  
ID AAP60599 standard; protein; 168 AA.

AC AAP60599;

DT 01-JAN-1980 (first entry)

DE Clone lambda HEPOL16 encoding human erythropoietin.

KM Erythropoietin; lambda HEPOL16; recombinant plasmid vector; anaemia;

KW mammal cell culture; 3T3; CHO; Chinese hamster ovary; ss.

OS Homo sapiens.

PN W06603520-A.

19-JUN-1986.

03-DEC-1985; 85WO-US02405.

22-JAN-1985; 85US-0693258.

04-DEC-1984; 84US-0677813.

03-JAN-1985; 85US-0688622.

(GENE-) GENETICS INST INC.

Fri Jan 3 10:09:19 2003

us-09-853-731-2.rag

Page 10

|   |   |   |
|---|---|---|
| P | A | (FRIT// FRITSCH E.  |
| X | X |   |
| P | I | Fritsch E, Hewick RM, Jacobs K;                                       |
| D | R | WPI: 1986-169459/26.  |
| X | X | N-PDDB; AAN60519.   |
| X | X |   |
| P | T | Prodn. of human cDNA clone expressing erythropoietin - for mass       |
| X | X | prodn. of erythropoietin, useful for treating anaemia                 |
| X | X |   |
| P | S | Disclosure; Page 20; 61pp; English.                                   |
| C | C |   |
| C | C | A recombinant plasmid vector expressing this clone is expressed in e. |
| C | C | g 3f3 or CHO cell cultures. The produced erythropoietin is useful     |
| C | C | for treatment of anaemia, especially renal anaemia. The cloned gene   |
| C | C | expresses high levels of the protein and thus provides a means of     |
| C | C | mass production. See also AAN60513-18, AAN60520-21 and AAP60598.      |
| X | X |   |
| S | Q | Sequence 188 AA:  |
|   |   |   |
| Q | U | Query Match 100.0%; Score 851; DB 7; Length 188;                      |
| B | E | Best Local Similarity 100.0%; Pred. No. 5.1e-87;                      |
| M | a | Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;          |
|   |   |   |
| Y | 1 | APPRICDSRVLYERYLLEAKENATTTCGAERCSINENTIVPDKTVNFMKRMVEVGOA 60          |
| D | b | 23 APPRICDSRVLYERYLLEAKENATTTCGAERCSINENTIVPDKTVNFMKRMVEVGOA 82       |
| O | y | 61 VEWOGGLALLSEAVLRGQALLVNSSQPWEPEQLHVDKAVSGIRSTTTLRALGAOKFAIS 120    |
| D | b | 83 VEWOGGLALLSEAVLRGQALLVNSSQPWEPEQLHVDKAVSGIRSTTTLRALGAOKFAIS 142    |
| O | y | 121 PPDAASAAPLRTTADTFRRKLERYVSINFRLRKLKYTGACRTGDR 166                 |
| D | b | 143 PPDAASAAPLRTTADTFRRKLERYVSINFRLRKLKYTGACRTGDR 188                 |

Search completed: January 2, 2003, 15:11:28  
Job time : 33.0967 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:10:20 ; Search time 8.52568 Seconds  
(without alignments)  
368.997 Million cell updates/sec

Title: US-09-853-731-2

Perfect score: 851

Sequence: 1 APPRLICDSRVLEKRYLLEAK.....NFLRGKIKLYTGACRTGDR 166

Scoring table: BLOSUM62

Searched: Gap0 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 851   | 100.0       | 166    | 10    | US-09-853-731-2     |
| 2          | 851   | 100.0       | 166    | 12    | US-10-014-363-2     |
| 3          | 851   | 100.0       | 169    | 12    | US-10-014-363-4     |
| 4          | 851   | 100.0       | 174    | 12    | US-10-014-363-3     |
| 5          | 851   | 100.0       | 174    | 12    | US-10-014-363-5     |
| 6          | 846   | 99.4        | 165    | 10    | US-09-853-731-1     |
| 7          | 846   | 99.4        | 165    | 12    | US-10-014-363-1     |
| 8          | 294   | 34.5        | 60     | 10    | US-09-864-761-48502 |
| 9          | 111   | 13.0        | 30     | 10    | US-09-975-063-2     |
| 10         | 111   | 13.0        | 30     | 12    | US-10-011-858-2     |
| 11         | 108   | 12.7        | 20     | 9     | US-09-919-703-9     |
| 12         | 88    | 10.3        | 20     | 9     | US-09-919-703-10    |
| 13         | 74    | 8.7         | 1564   | 10    | US-09-801-368-244   |
| 14         | 72.5  | 8.5         | 81     | 10    | US-09-764-877-1947  |
| 15         | 72.5  | 8.5         | 321    | 9     | US-09-945-182-26    |
| 16         | 70.5  | 8.3         | 813    | 9     | US-09-964-899-25    |
| 17         | 70    | 8.2         | 409    | 10    | US-09-815-242-14050 |
| 18         | 69.5  | 8.2         | 542    | 9     | US-10-045-815-8     |
| 19         | 69.5  | 8.2         | 747    | 9     | US-10-045-815-6     |

|    |      |     |      |    |                     |                   |
|----|------|-----|------|----|---------------------|-------------------|
| 20 | 69.5 | 8.2 | 1327 | 10 | US-09-841-835-2     | Sequence 2, App11 |
| 21 | 69.5 | 8.2 | 2472 | 10 | US-09-815-242-5064  | Sequence 5064, Ap |
| 22 | 68.5 | 8.0 | 327  | 9  | US-09-916-494A-16   | Sequence 16, App1 |
| 23 | 68   | 8.0 | 2201 | 12 | US-10-029-907-3     | Sequence 3, App1  |
| 24 | 67.5 | 7.9 | 236  | 10 | US-09-905-129-12    | Sequence 12, App1 |
| 25 | 67.5 | 7.9 | 236  | 10 | US-09-905-129-15    | Sequence 15, App1 |
| 26 | 67.5 | 7.9 | 236  | 10 | US-09-991-630-12    | Sequence 12, App1 |
| 27 | 67.5 | 7.9 | 236  | 10 | US-09-991-630-15    | Sequence 15, App1 |
| 28 | 67.5 | 7.9 | 379  | 9  | US-09-975-139-8     | Sequence 8, App1  |
| 29 | 67.5 | 7.9 | 766  | 10 | US-09-925-301-1276  | Sequence 1276, Ap |
| 30 | 67   | 7.9 | 136  | 10 | US-09-764-877-1623  | Sequence 1623, Ap |
| 31 | 67   | 7.9 | 1148 | 10 | US-09-815-242-13826 | Sequence 13826, A |
| 32 | 66.5 | 7.8 | 210  | 10 | US-09-738-626-6842  | Sequence 6842, Ap |
| 33 | 66.5 | 7.8 | 210  | 10 | US-09-731-872-387   | Sequence 387, App |
| 34 | 66   | 7.8 | 319  | 9  | US-09-815-242-10267 | Sequence 10267, A |
| 35 | 66   | 7.8 | 1618 | 10 | US-09-963-875-1     | Sequence 1, App1  |
| 36 | 66   | 7.8 | 7257 | 9  | US-10-014-717-5     | Sequence 5, App1  |
| 37 | 65.5 | 7.7 | 691  | 10 | US-09-815-242-12339 | Sequence 12339, A |
| 38 | 65.5 | 7.7 | 711  | 9  | US-09-738-626-3507  | Sequence 3507, Ap |
| 39 | 65   | 7.6 | 336  | 10 | US-09-925-300-1408  | Sequence 1408, Ap |
| 40 | 65   | 7.6 | 339  | 9  | US-09-961-876-246   | Sequence 246, App |
| 41 | 64   | 7.5 | 954  | 9  | US-09-944-413-7     | Sequence 7, App1  |
| 42 | 64   | 7.5 | 954  | 9  | US-09-944-403-7     | Sequence 7, App1  |
| 43 | 64   | 7.5 | 954  | 9  | US-09-944-896-7     | Sequence 7, App1  |
| 44 | 64   | 7.5 | 954  | 9  | US-09-944-944-7     | Sequence 7, App1  |
| 45 | 64   | 7.5 | 954  | 9  | US-09-944-907-7     | Sequence 7, App1  |

ALIGNMENTS

RESULT 1

US-09-853-731-2

: Sequence 2, Application US/09853731

: Patent No. US20020037841A1

: GENERAL INFORMATION:

: APPLICANT: Papadimitriou, Apollon

: TITLE OF INVENTION: Erythropoietin Composition

: FILE REFERENCE: 20619 US

: CURRENT APPLICATION NUMBER: US/09/853, 731

: CURRENT FILING DATE: 2001-05-11

: PRIOR APPLICATION NUMBER: EP/00010355.5

: PRIOR FILING DATE: 2000-05-15

: NUMBER OF SEQ ID NOS: 2

: SOFTWARE: PatentIn version 3.0

: SEQ ID NO 2

: LENGTH: 166

: TYPE: PRT

: ORGANISM: Homo sapiens

: US-09-853-731-2

Query Match 100.0%; Score 851; DB 10; Length 166;

Best Local Similarity 100.0%; Pred. No. 1, 1e-84;

Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEKRYLLEAKKAENITTCACHCISINENITVPDKVNFYAMKREVGQQA 60

DB 1 APPRLICDSRVLEKRYLLEAKKAENITTCACHCISINENITVPDKVNFYAMKREVGQQA 60

QY 61 VEWMOGALLSEAVYRGQALLVNSQPEPQLOLHVDKAVSGIRSLTTLRLGAKKEAIS 120

DB 61 VEWMOGALLSEAVYRGQALLVNSQPEPQLOLHVDKAVSGIRSLTTLRLGAKKEAIS 120

QY 121 PPDAASAPLRTITADTFPRKLFPRVSNFLRGKIKLYTGACRTGDR 166

DB 121 PPDAASAPLRTITADTFPRKLFPRVSNFLRGKIKLYTGACRTGDR 166

RESULT 2

US-10-014-363-2

: Sequence 2, Application US/10014363

: Patent No. US20020115833A1

: GENERAL INFORMATION:

```

; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; SOFTWARE: Patentin version 3.1
; NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 2
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-363-2

Query Match
Best Local Similarity 100.0%; Score 851; DB 12; Length 166;
Pred. No. 1,1e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRMEVGQA 60
D 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRMEVGQA 60
QY 61 VEWOGIALISEAVLRGOALLVNSQWPPEPLQIHVDKAVSGLSITLLRALGAOKRAIS 120
D 61 VEWOGIALISEAVLRGOALLVNSQWPPEPLQIHVDKAVSGLSITLLRALGAOKRAIS 120
QY 121 PPDAASAPLRTITADTFRKLFRVYSNPLRGKCLKLYTGEACRTGDR 166
D 121 PPDAASAPLRTITADTFRKLFRVYSNPLRGKCLKLYTGEACRTGDR 166
```

```

RESULT 3
US-10-014-363-4
; Sequence 4, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 169
; TYPE: PRT
; ORGANISM: CHO/dhfr-
US-10-014-363-4

Query Match
Best Local Similarity 100.0%; Score 851; DB 12; Length 169;
Pred. No. 1,1e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRMEVGQA 60
D 4 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRMEVGQA 63
QY 61 VEWOGIALISEAVLRGOALLVNSQWPPEPLQIHVDKAVSGLSITLLRALGAOKRAIS 120
D 64 VEWOGIALISEAVLRGOALLVNSQWPPEPLQIHVDKAVSGLSITLLRALGAOKRAIS 123
QY 121 PPDAASAPLRTITADTFRKLFRVYSNPLRGKCLKLYTGEACRTGDR 166
```

```

D 124 PPDAASAPLRTITADTFRKLFRVYSNPLRGKCLKLYTGEACRTGDR 169
```

```

RESULT 4
US-10-014-363-3
; Sequence 3, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 174
; TYPE: PRT
; ORGANISM: CHO/dhfr-
US-10-014-363-3
```

```

Query Match
Best Local Similarity 100.0%; Score 851; DB 12; Length 174;
Pred. No. 1,2e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRMEVGQA 60
D 9 APPRLICDSRYLERLYLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRMEVGQA 68
QY 61 VEWOGIALISEAVLRGOALLVNSQWPPEPLQIHVDKAVSGLSITLLRALGAOKRAIS 120
D 69 VEWOGIALISEAVLRGOALLVNSQWPPEPLQIHVDKAVSGLSITLLRALGAOKRAIS 128
QY 121 PPDAASAPLRTITADTFRKLFRVYSNPLRGKCLKLYTGEACRTGDR 166
D 129 PPDAASAPLRTITADTFRKLFRVYSNPLRGKCLKLYTGEACRTGDR 174
```

```

RESULT 5
US-10-014-363-5
; Sequence 5, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 174
; TYPE: PRT
; ORGANISM: CHO/dhfr-
US-10-014-363-5
```

```

Query Match
Best Local Similarity 100.0%; Score 851; DB 12; Length 174;
Pred. No. 1,2e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF053356.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
; OTHER INFORMATION: SWISSPROT HIT: P01588, EVALU8 9.00e-28
; OTHER INFORMATION: EST_HUMAN HIT: AA662379.1, EVALU8 3.00e-10
US-09-864-761-48502

Query Match          34.5%; Score 294; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.8e-25;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 VGGQAVEWQGLALSEAVIRGQALLVNSQPEPIQLHVDKAVSGIRSLTTLRALGAQ 115
DB 1 VGGQAVEWQGLALSEAVIRGQALLVNSQPEPIQLHVDKAVSGIRSLTTLRALGAQ 60

RESULT 9
US-09-975-063-2
Sequence 2, Application US/09975063
Patent No. US2002004525A1
GENERAL INFORMATION:
APPLICANT: POWELL, Jerry S.
TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL
NUMBER OF SEQUENCES: EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
CORRESPONDENCE ADDRESS:
ADDRESSER: DORSEY & WHITNEY LLP
STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.77/8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/975,063
FILING DATE: 10-Oct-2001
ATTORNEY/AGENT INFORMATION:
NAME: Robert's, Mark W.
REGISTRATION NUMBER: 46,160
REFERENCE/DOCKET NUMBER: 500582.03 (112893.109)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-903-8728
TELEFAX: 206-903-8820
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-975-063-2

Query Match          13.0%; Score 111; DB 10; Length 30;
Best Local Similarity 83.3%; Pred. No. 8.4e-06;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 APRRLICDSRYLERLYLLEAKEENITTCGA 30
DB 1 APRRLIDSRVLERLYLLEAKEEXITTDGCA 30

RESULT 10
US-10-011-858-2
Sequence 2, Application US/10011858
Patent No. US20020137145A1
GENERAL INFORMATION:
APPLICANT: POWELL, Jerry S.
```

```
; TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL.
; EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: DORSEY & WHITNEY, LLP
; STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS-WORD
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/011,858
; FILING DATE: 05-No. US20020137145A1-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: Robert's, Mark W.
; REGISTRATION NUMBER: 46,160
; REFERENCE/DOCKET NUMBER: 500582.13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-903-8728
; TELEFAX: 206-903-8820
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-011-858-2

Query Match          13.0%; Score 111; DB 12; Length 30;
Best Local Similarity 83.3%; Pred. No. 8.4e-06;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 APRRLICDSRYLERLYLLEAKEENITTCGA 30
DB 1 APRRLIDSRVLERLYLLEAKEEXITTDGCA 30

RESULT 11
US-09-919-703-9
Sequence 9, Application US/09919703
Patent No. US20020165129A1
GENERAL INFORMATION:
APPLICANT: Krystal, Gerald
APPLICANT: Radkin, Simon W.
TITLE OF INVENTION: Peptides and Their Use to Ameliorate
TITLE OF INVENTION: Cell Death
CURRENT APPLICATION NUMBER: US/09/919,703
FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 1998-04-59
PRIOR FILING DATE: 1998-04-59
PRIOR APPLICATION NUMBER: US 08/759,599
PRIOR FILING DATE: 1996-12-05
PRIOR APPLICATION NUMBER: US 60/008,233
PRIOR FILING DATE: 1995-12-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 9
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-9

Query Match          12.7%; Score 108; DB 9; Length 20;
```





Fri Jan 3 10:09:19 2003

us-09-853-731-2.rapb

Page 6

```

;
; COUNTRY: USA
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/945,182
; FILING DATE: 31-Aug-2001
; CLASSIFICATION: <unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,324
; FILING DATE: <unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
;
; INFORMATION FOR SEQ ID NO: 26:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-945-182-26

Query Match      8.5%; Score 72; DB 9; Length 321;
Best Local Similarity 27.2%; Pred. No. 3.8;
Matches 31; Conservative 11; Mismatches 48; Indels 24; Gaps 3;

QY 57 GQAQVEVWQGLALSEAVLRGQALLVNSQPEPLQIHDKAVSGLSLTLLRAIGAOK 116
   | : |||||
Db 77 GWEVFDVWQGL-----RHQFWKQLCLELRANGELDAGEAERAKRQPOD 120
   ||| |
QY 117 EAISPPDAASAAPLRTTADTFRKLFVYSGNFLRGKL-----KIYGEACRTG 164
   ||| |
Db 121 P--PPDLRSLGFGGRVRVPQPERALIVFTRSQRKNLFAEMRQLGSAAEAAGPG 172
```

Search completed: January 2, 2003, 15:13:50  
Job time : 9.52368 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:10:05 ; Search time 14.0423 Seconds  
(without alignments)  
1136.446 Million cell updates/sec

Title: US-09-853-731-2

Perfect score: 851

Sequence: 1 APPRLIDSRLEKRYLEAK.....NFLNGKLTLYGACRCRQDR 166

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.73:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 851   | 100.0       | 193    | 1 ZOHU   | erythropoietin pre |
| 2          | 769.5 | 90.4        | 192    | 1 J00173 | erythropoietin pre |
| 3          | 764.5 | 89.8        | 192    | 1 I84613 | erythropoietin pre |
| 4          | 718   | 84.4        | 188    | 1 I46083 | erythropoietin pre |
| 5          | 706   | 83.0        | 192    | 1 S28148 | erythropoietin pre |
| 6          | 690.5 | 81.1        | 194    | 1 I46401 | erythropoietin pre |
| 7          | 686   | 80.6        | 192    | 1 A24902 | erythropoietin pre |
| 8          | 685.5 | 80.6        | 195    | 1 JC7699 | erythropoietin - r |
| 9          | 683   | 80.3        | 190    | 2 I46578 | erythropoietin - p |
| 10         | 638   | 75.0        | 175    | 2 I46199 | erythropoietin - d |
| 11         | 90    | 10.6        | 353    | 2 G02729 | thrombopoietin - h |
| 12         | 89    | 10.5        | 353    | 2 I80105 | thrombopoietin pre |
| 13         | 88    | 10.3        | 323    | 2 AB0323 | erythropoietin pre |
| 14         | 87.5  | 10.3        | 346    | 2 AE0959 | ribonucleoside-dip |
| 15         | 86    | 10.1        | 286    | 2 A55530 | solute binding rec |
| 16         | 85    | 10.0        | 339    | 2 A83274 | megakaryocyte grow |
| 17         | 83.5  | 9.8         | 296    | 2 A10443 | UDP-N-acetylpyruv  |
| 18         | 80.5  | 9.5         | 3033   | 1 GNMVJ8 | probable 2-hydroxy |
| 19         | 79.5  | 9.3         | 3033   | 1 GNMVJ8 | genome polypeptide |
| 20         | 79    | 9.3         | 1829   | 2 T35681 | probable sensory h |
| 21         | 78.5  | 9.2         | 813    | 2 A56539 | ribosomal protein  |
| 22         | 78.5  | 9.2         | 897    | 2 AF0526 | ATP-dependent heli |
| 23         | 78    | 9.2         | 348    | 2 A54966 | EGF receptor subst |
| 24         | 78    | 9.2         | 455    | 2 T35450 | ABC transporter AT |
| 25         | 78    | 9.2         | 455    | 2 H97693 | methyamine utiliz  |
| 26         | 78    | 9.2         | 747    | 2 AG2919 | conserved hypothet |
| 27         | 77.5  | 9.1         | 242    | 1 S36741 | probable copper-tr |
| 28         | 77    | 9.0         | 451    | 2 AD1928 | hypothetical prote |
| 29         | 77    | 9.0         | 548    | 2 S75569 | hypothetical prote |
|            |       |             |        |          | symbolin synt. - p |

|    |      |     |      |          |                    |
|----|------|-----|------|----------|--------------------|
| 30 | 77   | 9.0 | 548  | 2 B84932 | 60 kD chaperonin [ |
| 31 | 76.5 | 9.0 | 154  | 2 H82810 | bacterioferritin X |
| 32 | 76.5 | 9.0 | 425  | 2 AE3465 | mandelate racemase |
| 33 | 75.5 | 8.9 | 544  | 2 S37039 | groEL protein - Ba |
| 34 | 75.5 | 8.9 | 637  | 2 S75772 | hypothetical prote |
| 35 | 74.5 | 8.8 | 400  | 2 AB2922 | conserved hypothet |
| 36 | 74.5 | 8.8 | 425  | 2 C97696 | ITS beta (AF305057 |
| 37 | 74.5 | 8.8 | 824  | 2 D64738 | ATP-dependent heli |
| 38 | 74   | 8.7 | 282  | 2 B37994 | RF2 protein - salm |
| 39 | 74   | 8.7 | 326  | 2 JC4125 | thrombopoietin pre |
| 40 | 74   | 8.7 | 335  | 2 AH3625 | ribonucleoside-dip |
| 41 | 74   | 8.7 | 544  | 2 B82048 | chaperonin, 60 kD  |
| 42 | 74   | 8.7 | 552  | 2 S39765 | chaperonin 60 - Co |
| 43 | 74   | 8.7 | 1089 | 2 S53978 | PSI protein - yea  |
| 44 | 74   | 8.7 | 1564 | 2 S55517 | probable transport |
| 45 | 73.5 | 8.6 | 401  | 2 H83911 | hypothetical prote |

#### ALIGNMENTS

RESULT 1  
ZOHU  
erythropoietin precursor [validated] - human  
C:Species: Homo sapiens (man)  
C:Date: 27-Nov-1985 #sequence revision 27-Nov-1985 #text\_change 08-Dec-2000  
C:Accession: A01855; A24744; A23284; A22210; S56178  
R:Jacobs, K.; Shoemaker, C.; Ruderstorff, R.; Neill, S.D.; Kaufman, R.J.; Mufson, A.;  
Nature 313, 806-810, 1985  
A>Title: Isolation and characterization of genomic and cDNA clones of human erythro  
A:Reference number: A01855; MUID:85137899; PMID:3838366  
A:Accession: A01855  
A:Molecule type: mRNA; DNA  
A:Residues: 1-193 <JAC>  
A:Cross-references: GB:X02157; GB:X02158  
R:Lin, F.K.; Sung, S.; Lin, C.H.; Browne, J.K.; Smalling, R.; Eyrle, J.C.; Chen, K.K  
Proc. Natl. Acad. Sci. U.S.A. 82, 7580-7584, 1985  
A>Title: Cloning and expression of the human erythropoietin gene.  
A:Reference number: A24744; MUID:86067948; PMID:3865178  
A:Accession: A24744  
A:Molecule type: DNA  
A:Residues: 1-193 <LIN>  
A:Cross-references: GB:M11319; NID:9182197; PIDN:AA52400.1; PID:9182198  
R:Li, P.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser, E.  
J. Biol. Chem. 261, 3116-3121, 1986  
A>Title: Structural characterization of human erythropoietin.  
A:Reference number: A25384; MUID:86140080; PMID:3949763  
A:Accession: A25384  
A:Molecule type: protein  
A:Residues: 28-86, 'Q', 87-193 <LAI>  
A:Experimental source: urine  
A:Note: forms without the carboxyl-terminal residue and the four carboxyl-terminal re  
R:Yanagawa, S.; Hirade, K.; Ohnoka, H.; Sasaki, R.; Chiba, H.; Ueda, M.; Goto, M.  
J. Biol. Chem. 259, 2707-2710, 1984  
A>Title: Isolation of human erythropoietin with monoclonal antibodies.  
A:Reference number: A22210; MUID:84135751; PMID:6698989  
A:Accession: A22210  
A:Molecule type: protein  
A:Residues: 28-29, 'X', 31-33, 'L', 35-50, 'X', 52-53, 'D', 55, 'G', 57 <YAN>  
R:Matsumoto, S.; Ikura, K.; Ueda, M.; Sasaki, R.  
Plant Mol. Biol. 27, 1163-1172, 1995  
A>Title: Characterization of a human glycoprotein (erythropoietin) produced in cultur  
A:Reference number: S56178; MUID:95284365; PMID:7766897  
A:Accession: S56178  
A:Molecule type: protein  
A:Residues: 28-33, 'X', 35-37 <MTS>  
A:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live  
C:Genetics: GDB:EPD  
A:Gene: GDB:EPD  
A:Cross-references: GDB:119110; OMIM:133170  
A:Map position: 7q21.3-7q22.1  
A:introns: 5/1; 53/3; 82/3; 142/3  
C:Function:

A:Description: the primary inducer of erythrocyte formation  
C:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver  
F:1-27/Domain: signal sequence #status predicted <Sig>  
F:28-193/Product: erythropoietin #status experimental <MAT>  
F:34-188,56-60/Disulfide bonds: #status experimental  
F:51,65,110/Binding site: carbohydrate (asn) (covalent) #status experimental  
F:153/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match  
Best Local Similarity 100.0%; Score 851; DB 1; Length 193;  
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLEAKKAENITTCGAHCSLNENITVPDTKVFYAKRMKEVGQQA 60  
DB 28 APPRLICDSRVLEERYLLEAKKAENITTCGAHCSLNENITVPDTKVFYAKRMKEVGQQA 87  
OY 61 VEWOGIALLSEAVLRQOALLVNSSQPEPQLHVDKAVGSLSTLTLLRALGAOKEAIS 120  
DB 88 VEWOGIALLSEAVLRQOALLVNSSQPEPQLHVDKAVGSLSTLTLLRALGAOKEAIS 147  
OY 121 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRTGDR 166  
DB 148 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRTGDR 193

RESULT 2  
JQ0173  
erythropoietin precursor - crab-eating macaque  
C:Species: Macaca fascicularis (crab-eating macaque)  
C>Date: 07-Sep-1990 #sequence\_revision 15-Nov-1996 #text\_change 22-Jun-1999  
C:Accession: JQ0173  
R:Lin, F.K.; Lin, C.H.; Lai, P.H.; Browne, J.K.; Egly, J.C.; Smalling, R.; Fox, G.M.;  
Gene 44, 201-209, 1986  
A:Title: Monkey erythropoietin gene: cloning, expression and comparison with the human  
A:Reference number: JQ0173; MUID:87055236; PMID:2877922  
A:Accession: JQ0173  
A:Molecule type: mRNA  
A:Residues: 1-192 <LIN>  
A:Cross-references: GB:M18189; GB:M15818; GB:M15819; GB:M18188; NID:9342093; PIDN:AAA368  
A:Experimental source: kidney  
C:Comment: This protein is the principal hormone involved in the regulation of erythrocy  
C:Function: Erythropoietin is produced by kidney or liver of adult mammals and by liver  
A:Description: the primary inducer of erythrocyte formation  
C:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver  
F:1-27/Domain: signal sequence #status predicted <Sig>  
F:28-192/Product: erythropoietin #status predicted <MAT>  
F:34-187,56-60/Disulfide bonds: #status predicted  
F:51,65,110/Binding site: carbohydrate (asn) (covalent) #status predicted  
F:152/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match  
Best Local Similarity 90.4%; Score 769.5; DB 1; Length 192;  
Matches 152; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLEAKKAENITTCGAHCSLNENITVPDTKVFYAKRMKEVGQQA 60  
DB 28 APPRLICDSRVLEERYLLEAKKAENITTCGAHCSLNENITVPDTKVFYAKRMKEVGQQA 87  
OY 61 VEWOGIALLSEAVLRQOALLVNSSQPEPQLHVDKAVGSLSTLTLLRALGAOKEAIS 120  
DB 88 VEWOGIALLSEAVLRQOALLVNSSQPEPQLHVDKAVGSLSTLTLLRALGAOKEAIS 146  
OY 121 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRTGDR 166  
DB 147 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRTGDR 192

RESULT 3  
184613  
erythropoietin precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)  
C>Date: 02-Aug-1996 #sequence\_revision 15-Nov-1996 #text\_change 22-Jun-1999  
C:Accession: I84613  
R:Men, D.; Boissel, J.  
Blood 82, 1507-1516, 1993  
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom  
A:Reference number: I46083; MUID:93372347; PMID:8364201  
A:Accession: I84613  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-192 <RES>  
A:Cross-references: GB:L10609; NID:9342095; PIDN:AAA36842.1; PID:9342096  
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live  
C:Function: Erythropoietin is produced by kidney or liver of adult mammals and by live  
A:Description: the primary inducer of erythrocyte formation  
C:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver  
F:1-27/Domain: signal sequence #status predicted <Sig>  
F:28-192/Product: erythropoietin #status predicted <MAT>  
F:34-187,56-60/Disulfide bonds: #status predicted  
F:51,65,110/Binding site: carbohydrate (asn) (covalent) #status predicted  
F:152/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match  
Best Local Similarity 89.8%; Score 764.5; DB 1; Length 192;  
Matches 150; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLEAKKAENITTCGAHCSLNENITVPDTKVFYAKRMKEVGQQA 60  
DB 28 APPRLICDSRVLEERYLLEAKKAENITTCGAHCSLNENITVPDTKVFYAKRMKEVGQQA 87  
OY 61 VEWOGIALLSEAVLRQOALLVNSSQPEPQLHVDKAVGSLSTLTLLRALGAOKEAIS 120  
DB 88 VEWOGIALLSEAVLRQOALLVNSSQPEPQLHVDKAVGSLSTLTLLRALGAOKEAIS 146  
OY 121 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRTGDR 166  
DB 147 PPDASAAPLRTITADTFERKLFYRYSNPLRGKLTLYTGECRCRTGDR 192

RESULT 4  
I46083  
erythropoietin precursor - cat (fragment)  
C:Species: Felis silvestris catus (domestic cat)  
C>Date: 16-Aug-1996 #sequence\_revision 15-Nov-1996 #text\_change 22-Jun-1999  
C:Accession: I46083  
R:Men, D.; Boissel, J.  
Blood 82, 1507-1516, 1993  
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom  
A:Reference number: I46083; MUID:93372347; PMID:8364201  
A:Accession: I46083  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-188 <MEN>  
A:Cross-references: GB:L10606; NID:9163820; PIDN:AAA30807.1; PID:9163821  
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live  
C:Function: Erythropoietin is produced by kidney or liver of adult mammals and by live  
A:Description: the primary inducer of erythrocyte formation  
C:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver  
F:1-27/Domain: signal sequence (fragment) #status predicted <Sig>  
F:23-188/Product: erythropoietin #status predicted <MAT>  
F:29-183,51-55/Disulfide bonds: #status predicted  
F:46,60,105/Binding site: carbohydrate (asn) (covalent) #status predicted  
F:148/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match  
Best Local Similarity 84.4%; Score 718; DB 1; Length 188;  
Matches 140; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLEAKKAENITTCGAHCSLNENITVPDTKVFYAKRMKEVGQQA 60  
DB 23 APPRLICDSRVLEERYLLEAKKAENITTCGAHCSLNENITVPDTKVFYAKRMKEVGQQA 82

QY 61 VEEQGLALISEAVLRGQALLVNSGQWELQIHYKVAAGLSITTLTALGAKQEAIS 120  
 |||||  
 Db 83 VEEVQGLALISEALLRGQALLANSQSEITLQIHLVDAVASLSISITSLPALGAKQEAIS 142  
 |||||  
 QY 121 PPDAAAPLRTITADTFKRLFRVYSNFLRGKLLKLYTGEACRTGDR 166  
 |||||  
 Db 143 LPEVTSAPLRTFTVDTLCKLFRVYSNFLRGKLLKLYTGEACRTGDR 188  
 |||||

RESULT 5  
S28148  
erythropoietin precursor - rat

|                       |       |              |     |            |       |        |     |
|-----------------------|-------|--------------|-----|------------|-------|--------|-----|
| Query Match           | 83.0% | Score        | 706 | DB         | 1     | Length | 192 |
| Best Local Similarity | 82.5% | Pred.        | No. | 7          | 8e-61 |        |     |
| Matches               | 137   | Conservative | 13  | Mismatches | 16    | Indels | 0   |
|                       |       |              |     |            |       | Gaps   | 0   |

RESULT 6  
 I46401  
 erythropoietin precursor - sheep  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 16-Aug-1996 #sequence\_revision 15-Nov-1996 #text\_change 22-Jun-1999  
 C:Accession: I46401; I47077  
 R:Fu, P.; Evans, B.; Lim, G.B.; Moritz, K.; Wintour, E.M.  
 Mol. Cell. Endocrinol. 93, 107-116, 1993  
 A:Title: The sheep erythropoietin gene: molecular cloning and effect of hemorrhage on p  
 A:Reference number: I46401; MUID:93351736, PMID:8349021  
 A:Accession: I46401  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA

A:Residues: 1,94 <PDB>  
A:Cross-references: EMBL:Z24681; NID:g395049; PIDD:CA80848.1; PID:g395050  
R:Wen, D.; Boissel, J.  
Blood 82, 1507-1516, 1993  
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom  
A:Reference number: I46083; MUID:93372347; PMID:8364201  
A:Accession: 147077  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 4-15, 'L', 17-107, 'P', 109-194 <MEN>  
A:Cross-references: GB:L10610; NID:g165876; PIDD:AAA31518.1; PID:g165877  
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live  
C:Function:  
A:Description: the primary inducer of erythrocyte formation  
C:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver  
F:1-37/Domain: signal sequence #status predicted <SIG>  
F:28-194/Product: erythropoietin #status predicted <MAT>  
F:34-189,56-60/Disulfide bonds: #status predicted  
F:51,65,110/Binding site: carboxylate (asn) (covalent) #status predicted  
F:154/Binding site: carboxylate (Ser) (covalent) #status predicted

Query Match 81.1%; Score 690.5; DB 1; Length 194;  
Best Local Similarity 82.0%; Pred. NO. 2.5e-59;  
Matches 137; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

|    | Query Match | Best Local Similarity                                       | Matched 137: Conservative | 81.18: Score 690.5: DB 1: Length 194: Pred. No. 2.5e-59: Mismatches 9: Indels 1: Gaps |
|----|-------------|---|---------------------------|---|
| Qy | 1           | APPRLICDRLVLEYLEAKEAEVNTTGCACACGSLNINITYPDTKVFVFAMKREMEYQQA | 60                        |   |
| Db | 28          | APPRLICDRLVLEYLEAKEAEVNTTGCACACGSLNINITYPDTKVFVFAMKREMEYQQA | 87                        |   |
| Qy | 61          | VEWVGALLSEAVLRGALLVNSSQPWEPLQIHHVDKAVSGLSRLTTLRALGAQKEAIS   | 120                       |   |
| Db | 88          | LEWVGALLSEAVLRGALLVNSSQPWEPLQIHHVDKAVSGLSRLTTLRALGAQKEAIS   | 147                       |   |
| Qy | 121         | PPDDA-SAAPRLTITADPFRKLFVSNFLRGKAKLYTGGACRGCDR               | 166                       |   |
| Db | 148         | LPDAPSAAPRLTITADPFRKLFVSNFLRGKAKLYTGGACRGCDR                | 194                       |   |

RESULT 7  
A24902  
erythropoietin precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 25-Oct-1987 #sequence\_revision 15-Nov-1996 #text\_change 22-Jun-1999  
C:Accession: A24902; A24901  
R:Shoemaker, C.B.; Miltsock, L.D.  
Mol. Cell. Biol. 6, 849-858, 1986  
A:Title: Murine erythropoietin gene: cloning, expression, and human gene homology.  
A:Reference number: A24902; MUID:87039103; PMID:3773894  
A:Accession: A24902  
A:Molecule type: DNA  
A:Residues: 1-192 <SHO>  
A:Note: the authors translated the codon TTA for residue 12 as Phe, TTA for residue 4  
R:McDonald, J.D.; Lin, F.K.; Goldwasser, E.  
Mol. Cell. Biol. 6, 842-848, 1986  
A:Title: Cloning, sequencing, and evolutionary analysis of the mouse erythropoietin g  
A:Reference number: A24901; MUID:87039104; PMID:3022133  
A:Accession: A24901  
A:Molecule type: DNA  
A:Residues: 1-67, 'P', 69-192 <MCD>  
A:Cross-references: GB:M2930; NID:g193086; PIND:AAA37570.1; PID:g387152  
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver  
C:Genetics:  
A:Introns: 5/1, 52/3; 81/3; 141/3  
C:Function:  
A:Description: the primary inducer of erythrocyte formation  
C:Superfamily: erythropoietin  
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver  
F:1-36/Domain: signal sequence #status predicted <SIG>  
F:27-192/Product: erythropoietin #status predicted <MAT>  
F:33-187, 55-165/Disulfide bonds: #status predicted  
F:50, 64, 109/Binding site: carbohydrate (asn) #status predicted

Query Match 80.6%; Score 686; DB 1; Length 192;  
Best Local Similarity 79.5%; Pred. No. 6,7e-59;  
Matches 132; Conservative 14; Mismatches 20; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLEAKENITTCGAHCISLNTNITVPDTKVNPFYAKRMVEVGQA 60  
DB 27 APPRLICDSRVLEERYLLEAKENITTCGAHCISLNTNITVPDTKVNPFYAKRMVEVGQA 86  
OY 61 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHYDKAVSGRLSTLTLLRALGAKKEAIS 120  
DB 87 LEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHYDKAVSGRLSTLTLLRALGAKKEAIS 146  
OY 121 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECACRTGDR 166  
DB 147 PPDTPPAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECACRTGDR 192

## RESULT 8

JC7699

erythropoietin - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001

C:Accession: JC7699

R:Vialalta, A.; Wu, D.; Margalith, M.; Hobart, P.

Biochem. Biophys. Res. Commun. 284, 823-827, 2001

A:Title: Rabbit Epo gene and cDNA: Expression of rabbit Epo after intramuscular injectio

A:Reference number: JC7699; MUID:21290682; PMID:11396976

A:Contents: kidney

A:Accession: JC7699

A:Molecule type: DNA

A:Residues: 1-195 &lt;VIL&gt;

A:Cross-references: GB:AF290943

C:Comment: This protein, a heavily glycosylated 34k protein produced in the fetal liver

cytes.

C:Genetics:

A:Gene: epo

C:Superfamily: erythropoietin

C:Keywords: glycoprotein; kidney

Query Match 80.6%; Score 685.5; DB 2; Length 195;  
Best Local Similarity 81.4%; Pred. No. 7,7e-59;  
Matches 136; Conservative 12; Mismatches 18; Indels 1; Gaps 1;

OY 1 APPRLICDSRVLEERYLLEAKENITTCGAHCISLNTNITVPDTKVNPFYAKRMVEVGQA 60  
DB 29 APPRLICDSRVLEERYLLEAKENITTCGAHCISLNTNITVPDTKVNPFYAKRMVEVGQA 88  
OY 61 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHYDKAVSGRLSTLTLLRALGAKKEAIS 120  
DB 89 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHYDKAVSGRLSTLTLLRALGAKKEAIS 148  
OY 121 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECACRTGDR 166  
DB 149 PPEAASAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECACRTGDR 195

## RESULT 9

I46578

erythropoietin - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 16-Jul-1999

C:Accession: I46578

R:Men, D.; Boissel, J.

Blood 82, 1507-1516, 1993

A:Title: Erythropoietin structure-function relationships: High degree of sequence homolo

A:Reference number: I46083; MUID:93372347; PMID:8364201

A:Accession: I46578

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-190 &lt;MEN&gt;

A:Cross-references: GB:I10607; NID:g164445; PIDN:AAA31029.1; PID:g164446

C:Superfamily: erythropoietin

Query Match 80.3%; Score 683; DB 2; Length 190;  
Best Local Similarity 82.1%; Pred. No. 1,3e-58;  
Matches 138; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

OY 1 APPRLICDSRVLEERYLLEAKENITTCGAHCISLNTNITVPDTKVNPFYAKRMVEVGQA 60  
DB 23 APPRLICDSRVLEERYLLEAKENITTCGAHCISLNTNITVPDTKVNPFYAKRMVEVGQA 82  
OY 61 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHYDKAVSGRLSTLTLLRALGAKKEAIS 120  
DB 83 MEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHYDKAVSGRLSTLTLLRALGAKKEAIS 142  
OY 121 PPDA--ASAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECACRTGDR 166  
DB 143 LPDASPSAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECACRTGDR 190

## RESULT 10

I46199

erythropoietin - dog (fragment)

C:Species: Canis lupus familiaris (dog)

C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 23-Jul-1999

C:Accession: I46199

R:Men, D.; Boissel, J.

Blood 82, 1507-1516, 1993

A:Title: Erythropoietin structure-function relationships: High degree of sequence hom

A:Reference number: I46083; MUID:93372347; PMID:8364201

A:Accession: I46199

A:Molecule type: mRNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-175 &lt;MEN&gt;

A:Cross-references: GB:I13027; NID:g290087; PIDN:AAA30842.1; PID:g552347

C:Superfamily: erythropoietin

Query Match 75.0%; Score 638; DB 2; Length 175;  
Best Local Similarity 81.0%; Pred. No. 2,6e-54;  
Matches 124; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLEERYLLEAKENITTCGAHCISLNTNITVPDTKVNPFYAKRMVEVGQA 60  
DB 23 APPRLICDSRVLEERYLLEAKENITTCGAHCISLNTNITVPDTKVNPFYAKRMVEVGQA 82  
OY 61 VEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHYDKAVSGRLSTLTLLRALGAKKEAIS 120  
DB 83 LEWMOGLALISEAVLRGQALLVNSSOPPEPQLQHYDKAVSGRLSTLTLLRALGAKKEAIS 142  
OY 121 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECACRTGDR 153  
DB 143 LPEASAPLRTITADTFKRLFRVYSNPLRGKLTLYTGECACRTGDR 175

## RESULT 11

G02729

thrombopoietin - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 05-Nov-1999

C:Accession: G02729

R:Im, S.

submitted to the EMBL Data Library, May 1996

A:Reference number: H01637

A:Accession: G02729

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-353 &lt;IMX&gt;

A:Cross-references: EMBL:U59493; NID:g1401245; PIDN:AAB03392.1; PID:g1401246

C:Genetics: hipo

Query Match 10.6%; Score 90; DB 2; Length 353;  
Best Local Similarity 26.3%; Pred. No. 0.61;  
Matches 41; Conservative 20; Mismatches 75; Indels 20; Gaps 5;

OY 1 APPRLICDSRVLEERYLLEAKENITTCGAHCISLNTNITVPDTKVNPFYAKRMVEVGQA 60

Db 24 APP--ACDLRLVLSKILRLSHVLSKLSQCEPVEHLPPLPVLLPAVDPSIGKEKTKTOMEETKA 81  
 Oy 61 VEWOGIALSEAVL--RGQALLVNSSQPEWPEIQLHVDKAVSGSLRSLTLLRALGAOKEA 118  
 Db 82 QDILGAVVLLLEGVMAARGQLGPTCLSSLLGQLSEQVRLILGALQSL-----LGTQ--- 132  
 Oy 119 ISPPMAASAPLRLTITADTFPKRLFRVRYSNFLRGKLK 154  
 Db 133 -LPPQG-----RTTAHKDPNAIFLSFQHLLRGKVR 161  
 RESULT 12  
 180105  
 thrombopoietin precursor - human  
 N:Alternate names: c-MPL ligand; megakaryocyte growth and development factor precursor  
 C:Species: Homo sapiens (man)  
 C:Date: 24-May-1996 #sequence, revision 24-May-1996 #text, change 20-Jun-2000  
 C:Accession: 159281; 180105; S45331; S48740; 138672; 152610  
 P:Poster, D.C.; Sprecher, C.A.; Grant, F.J.; Kramer, J.M.; Kijpfer, J.L.; Holly, R.D.; W  
 Proc. Natl. Acad. Sci. U.S.A. 91, 13023-13027, 1994  
 A>Title: Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosomal  
 A:Reference number: 159281; MUID:95108091; PMID:7809166  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-353 <RE2>  
 A:Cross-references: GB:L36051; NID:g533214; PIDN:AAAC37568.1; PID:g533215  
 A:Accession: 180105  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-353 <RE3>  
 A:Cross-references: GB:S36052; NID:g533216; PIDN:AAC37566.1; PID:g533217  
 R:De Sauvage, F.J.; Hass, P.E.; Spencer, S.D.; Malloy, B.E.; Gurney, A.L.; Spencer, S.A  
 D.V.; Eaton, D.L.  
 Nature 369, 533-538, 1994  
 A>Title: Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl ligand.  
 A:Reference number: S45331; MUID:94261202; PMID:8202154  
 A:Accession: S45331  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-353 <SNU>  
 A:Cross-references: GB:L33410; NID:g506826; PIDN:AAA59857.1; PID:g506827  
 R:Somma, Y.; Aahorri, H.; Seki, N.; Horl, T.; Ogami, K.; Kato, T.; Shinada, Y.; Kawamura  
 FEBS Lett. 353, 57-61, 1994  
 A>Title: Molecular cloning and chromosomal localization of the human thrombopoietin gene  
 A:Reference number: S48740; MUID:95010765; PMID:7926023  
 A:Accession: S48740  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-353 <SOH>  
 A:Cross-references: GB:D32046; NID:g577319; PIDN:BAAO6807.1; PID:g577320  
 R:Barclay, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.S.;  
 Cell 77, 1117-1124, 1994  
 A>Title: Identification and cloning of a megakaryocyte growth and development factor tha  
 A:Reference number: A54463; MUID:94291201; PMID:8020099  
 A:Accession: 138672  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-112, 'E', 114-353 <RE3>  
 A:Cross-references: EMBL:U11025; NID:g511223; PIDN:AAA50553.1; PID:g558078  
 R:Gurney, A.L.; Khang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.  
 Blood 85, 981-988, 1995  
 A>Title: Genomic structure, chromosomal localization, and conserved alternative splice f  
 A:Reference number: 152610; MUID:95152076; PMID:7849319  
 A:Accession: 152610  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-353 <RE4>  
 A:Cross-references: GB:S76771; NID:g914225; PIDN:AAB33390.1; PID:g914226  
 C:Genetics:  
 A:Gene: GDB:THPO; MGD#  
 A:Cross-references: GDB:374007; OMIM:600044

[illegible]

Fri Jan 3 10:09:20 2003

us-09-853-731-2.rpt

Page 6

A:Reference number: AB0502; PMID:11677608  
A:Accession: AE0959  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-346 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAMD03169.1; PID:J16504804; GSPDB:GN00176  
A:Genetics:  
A:Gene: STY3952

|                       |                 |               |           |            |
|-----------------------|-----------------|---------------|-----------|------------|
| Query Match           | 10.3%           | Score 87.5    | DB 2      | Length 346 |
| Best Local Similarity | 26.7%           | Ped No. 1     |           |            |
| Matches 44            | Conservative 22 | Mismatches 48 | Indels 51 | Gaps 9     |

  

|    |     |                        |                              |                            |     |
|----|-----|------------------------|------------------------------|----------------------------|-----|
| QY | 10  | RVERLYLLEKAEKENTTG-    | CAEHCISLNE                   | -NITVPTKYNFAMKRMVEGOOAVEWQ | 65  |
|    |     |                        |                              |                            |     |
| Db | 217 | RNLQLELHNPANVAYASAI    | IAEAKMEGCKNLTPLTIVSFL        | -----THQVYR                | 267 |
|    |     |                        |                              |                            |     |
| QY | 66  | GLALISAVLVGALLVNSQ     | -PWEPLDIHYDKAVSGLSRTLTLLRAGQ | -KEAISPP                   | 122 |
|    |     |                        |                              |                            |     |
| Db | 268 | GLK-----KGIILNALSDQAMQ | -----GLATIOSIKYLOQ           | PVPENISPP                  | 309 |
|    |     |                        |                              |                            |     |
| QY | 123 | -----DAASAPLRTTADPTFRK | LFVYVSNFLNGKLIKLYTGEA        | 160                        |     |
|    |     |                        |                              |                            |     |
| Db | 310 | VLLTTHNNADSARVRSLS     | SPGFRFVY-----LYOTYSPA        | 344                        |     |
|    |     |                        |                              |                            |     |

```

RESULT 15
A:Accession: A55530
M: megakaryocyte growth and development factor, long form - human
N: Alternate names: MPL, ligand, long form
C: Species: Homo sapiens (man)
C: Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text-change 07-May-1999
C: Accession: A55530
R: Chang, M.; McNinch, J.; Basu, R.; Shutter, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs,
J. Biol. Chem. 270, 511-514, 1995
A: Title: Cloning and characterization of the human megakaryocyte growth and development
A: Reference number: A55530; MUID:95122483; PMID:7822271
A: Accession: A55530
A: Status: Preliminary: not compared with conceptual translation
A: Molecule type: DNA
A: Residues: 1-286 <CHA>
A: Cross-references: GB:U17071
C: Genetics:
A: Gene: MGDF
A: Map position: 3q26.3
A: Keywords: alternative splicing; cytokine

```

| Query Match           | 10.1%  | Score 86;  | DB 2;          | Length 286;      |
|-----------------------|--------|--|----------------|------------------|
| Best Local Similarity | 26.6%; | Pred No. 1.2;  | Mismatches 41; | conservative 18; |
| Matches               | 41;    | conservative 18;   | Mismatches 41; | conservative 18; |
| QY                    | 1      | APPRLICISLREYLLLEAKENNTTGCACHECISINENTIVDTQVNFYAMRMEVGQA   | 60             |                  |
| Db                    | 24     | APP--ACDLRVSKILRSDHYHLSRSCQPEVHPILPTPLPAVDFSLGEMWTQMEETKA  | 81             |                  |
| QY                    | 61     | VWVWGGLALISAVL--RQALLVNSSQPEFLQIHYKRAVSGLSITLTLLRATLGAQKEA | 118            |                  |
| Db                    | 82     | ODILCAVTLLEGVMAAKRGPTGLSITLSLQSLGQVRLILGALQSL-----LSTQ---  | 132            |                  |
| QY                    | 119    | ISPPDAASAPLRTTTADTFERLLRFVYNSFLNGK                         | 152            |                  |
| Db                    | 133    | LPPQG-----RTTAHKDPNAIFLSFOHLRGK                            | 159            |                  |

Search completed: January 2, 2003, 15:13:27  
Job time : 15.0423 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2003, 15:07:09 ; Search time 8.52568 Seconds  
(without alignments)  
807.569 Million cell updates/sec

Title: US-09-853-731-2

Perfect score: 851  
Sequence: 1 APRRLIDSRVLEKRYLLEAK.....NFLRGKIKLYTGECRTGDR 166

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 851   | 100.0       | 193    | 1     | EPO_HUMAN          |
| 2          | 769.5 | 90.4        | 192    | 1     | EPO_MACFA          |
| 3          | 764.5 | 89.8        | 192    | 1     | EPO_MACMU          |
| 4          | 711   | 83.5        | 192    | 1     | EPO_PELCA          |
| 5          | 706   | 83.0        | 192    | 1     | EPO_RAT            |
| 6          | 697.5 | 82.0        | 192    | 1     | EPO_BOVIN          |
| 7          | 694   | 81.6        | 192    | 1     | EPO_MOUSE          |
| 8          | 690.5 | 81.1        | 194    | 1     | EPO_SHEEP          |
| 9          | 683   | 80.3        | 190    | 1     | EPO_PIG            |
| 10         | 638   | 75.0        | 175    | 1     | EPO_CANFA          |
| 11         | 109   | 12.8        | 352    | 1     | TPO_CANFA          |
| 12         | 89    | 10.5        | 353    | 1     | TPO_HUMAN          |
| 13         | 82    | 9.6         | 548    | 1     | CH60_BUCAP         |
| 14         | 81    | 9.5         | 551    | 1     | CH60_BUCAP         |
| 15         | 80.5  | 9.5         | 3033   | 1     | POLG_HCVJ8         |
| 16         | 79.5  | 9.3         | 552    | 1     | CH60_PSEST         |
| 17         | 78.5  | 9.2         | 897    | 1     | EP15_MOUSE         |
| 18         | 78    | 9.2         | 747    | 1     | ATCS_SYN7          |
| 19         | 77    | 9.0         | 548    | 1     | CH60_BUCAI         |
| 20         | 76.5  | 9.0         | 386    | 1     | CTBP_DROME         |
| 21         | 75.5  | 8.9         | 543    | 1     | CH60_BARBA         |
| 22         | 75    | 8.8         | 547    | 1     | CH60_LEGPN         |
| 23         | 74.5  | 8.8         | 809    | 1     | HRPB_ECOLI         |
| 24         | 74    | 8.7         | 326    | 1     | TPO_RAT            |
| 25         | 74    | 8.7         | 544    | 1     | CH60_FRATU         |
| 26         | 74    | 8.7         | 552    | 1     | CH60_COXBU         |
| 27         | 74    | 8.7         | 1089   | 1     | IMB3_YEAST         |
| 28         | 74    | 8.7         | 1564   | 1     | PDRA_YEAST         |
| 29         | 73.5  | 8.6         | 830    | 1     | GCL2_MOUSE         |
| 30         | 73.5  | 8.6         | 837    | 1     | GCL2_MOUSE         |
| 31         | 73    | 8.6         | 263    | 1     | GCL2_HUMAN         |
| 32         | 72.5  | 8.5         | 762    | 1     | YH25_DEIRA         |
| 33         | 72    | 8.5         | 356    | 1     | SLAP_ACEKI         |
|            |       |             |        |       | TPO_MOUSE          |
|            |       |             |        |       | P40226 mus musculu |

|    |      |     |      |   |            |
|----|------|-----|------|---|------------|
| 34 | 71   | 8.3 | 547  | 1 | CH60_PASMU |
| 35 | 70.5 | 8.3 | 217  | 1 | YND_ECOLI  |
| 36 | 70.5 | 8.3 | 381  | 1 | MODD_MYCAN |
| 37 | 70.5 | 8.3 | 896  | 1 | EP15_MOUSE |
| 38 | 70   | 8.2 | 319  | 1 | RIR4_SALTY |
| 39 | 70   | 8.2 | 544  | 1 | CH60_AERSA |
| 40 | 69.5 | 8.2 | 544  | 1 | CH60_METIO |
| 41 | 69.5 | 8.2 | 547  | 1 | CH60_PSEPU |
| 42 | 69.5 | 8.2 | 547  | 1 | CH60_PSEAU |
| 43 | 69.5 | 8.2 | 551  | 1 | CH60_AMORS |
| 44 | 69.5 | 8.2 | 907  | 1 | GACS_PSESY |
| 45 | 69.5 | 8.2 | 1327 | 1 | TNKL_HUMAN |

## ALIGNMENTS

|  |                        |
|--|------------------------|
| RESULT 1   |                        |
| EPO_HUMAN  | STANDARD: PRT: 193 AA. |
| AC P01588; Q9UHA0; Q9UE25; Q9UD20;                                       |                        |
| DR 21-JUL-1986 (Rel. 01, Created)  |                        |
| DR 21-JUL-1986 (Rel. 01, Last sequence update)                           |                        |
| DE 16-OCT-2001 (Rel. 40, Last annotation update)                         |                        |
| DE Erythropoietin precursor (Epoetin).                                   |                        |
| GN EPO.  |                        |
| OS Homo sapiens (Human).   |                        |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;     |                        |
| OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.             |                        |
| OX NCBI_TaxID=9606;  |                        |
| RN [1]   |                        |
| RP SEQUENCE FROM N.A.  |                        |
| RX MEDLINE=85137899; PubMed=3838366;                                     |                        |
| RA Jacobs K., Shoemaker C., Rudersdorf R., Neill S.D., Kaufman R.J.,     |                        |
| RA Mufson A., Seehra J., Jones S.S., Hewick R., Fritsch E.F.,            |                        |
| RA Kawakita M., Shimizu T., Miyake T.;                                   |                        |
| RT "Isolation and characterization of genomic and cDNA clones of human   |                        |
| RT erythropoietin."  |                        |
| RL Nature 313:806-810(1985).   |                        |
| RN [2]   |                        |
| RP SEQUENCE FROM N.A.  |                        |
| RX MEDLINE=86067948; PubMed=3865178;                                     |                        |
| RA Lin F.-K., Suggs S., Lin C.-H., Browne J.K., Smalling R., Egrie J.C., |                        |
| RA Tsui L.-C., Rosenthal A.;   |                        |
| RT "Large-scale sequencing of two regions in human chromosome 7q22:      |                        |
| RT analysis of 650 kb of genomic sequence around the EPO and CFHL1 loci  |                        |
| RT reveals 17 genes."  |                        |
| RL Genome Res. 8:1060-1073(1998).  |                        |
| RN [4]   |                        |
| RP SEQUENCE FROM N.A.  |                        |
| RA Rupert J.L., Hochachka P.W.;  |                        |
| RT "Erythropoietin gene sequence in the Quechua, a high altitude native  |                        |
| RT population."  |                        |
| RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.               |                        |
| RN [5]   |                        |
| RP SEQUENCE OF 58-193 FROM N.A., AND VARIANTS HEPATOCELLULAR CARCINOMA.  |                        |
| RX MEDLINE=93384593; PubMed=8396923;                                     |                        |
| RA Funakoshi A., Muta H., Baba T., Shimizu S.;                           |                        |
| RT "Gene expression of mutant erythropoietin in hepatocellular           |                        |
| RT carcinoma."   |                        |
| RL Biochem. Biophys. Res. Commun. 195:717-722(1993).                     |                        |
| RN [6]   |                        |
| RP SEQUENCE OF 28-193, AND DISULFIDE BONDS.                              |                        |
| RC TISSUE-Urine;   |                        |
| RX MEDLINE=86140080; PubMed=3949763;                                     |                        |

RA Lai P.H., Everett R., Wang F.F., Arakawa T., Goldwasser E.;  
 RT "Structural characterization of human erythropoietin.";  
 RL J. Biol. Chem. 261:3116-3121(1986).  
 [7]  
 RP PRELIMINARY SEQUENCE OF 28-57.  
 RX MEDLINE=84135751; PubMed=6698989;  
 RA Yanagawa S., Hirade K., Ohnota H., Sasaki R., Chiba H., Ueda M.,  
 Goto M.;  
 RT "Isolation of human erythropoietin with monoclonal antibodies.";  
 RL J. Biol. Chem. 259:2707-2710(1984).  
 [8]  
 RP STRUCTURE OF CARBOHYDRATES.  
 RX MEDLINE=88153657; PubMed=3346214;  
 RA Takeuchi M., Takasaki S., Miyazaki H., Kato T., Hoshi S., Kochibe N.,  
 Kobata A.;  
 RT "Comparative study of the asparagine-linked sugar chains of human  
 erythropoietins purified from urine and the culture medium of  
 recombinant Chinese hamster ovary cells.";  
 RL J. Biol. Chem. 263:3657-3663(1988).  
 [9]  
 RP STRUCTURE OF CARBOHYDRATES.  
 RX MEDLINE=89118279; PubMed=3219367;  
 RA Sasaki H., Ochi N., Dell A., Fukuda M.;  
 RT "Site-specific glycosylation of human recombinant erythropoietin:  
 analysis of glycopeptides or peptides at each glycosylation site by  
 fast atom bombardment mass spectrometry.";  
 RL Biochemistry 27:8618-8626(1988).  
 [10]  
 RP STRUCTURE OF CARBOHYDRATES.  
 RX MEDLINE=92314463; PubMed=1820196;  
 RA Takeuchi M., Kobata A.;  
 RT "Structures and functional roles of the sugar chains of human  
 erythropoietins.";  
 RL Glycobiology 1:337-346(1991).  
 [11]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=98445092; PubMed=9774108;  
 RA Syed R.S., Reid S.W., Li C., Cneethan J.C., Aoki K.H., Liu B.,  
 Zhen H., Ossling T.D., Chirino A.J., Zhang J., Finer-Moore J.,  
 Elliott S., Stoney K., Katz B.A., Matthews D.J., Wendoloski J.J.,  
 Egnell J., Strong K.M.;  
 RT "Efficiency of signaling through cytokine receptors depends  
 critically on receptor orientation.";  
 RL Nature 395:511-516(1998).  
 [12]  
 RP FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE  
 REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A  
 PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS  
 AND BY LIVER OF FETAL OR NEONATAL MAMMALS.  
 CC -1- PHARMACEUTICAL: Used for the treatment of anemia. Available under  
 the names Epogen (Amgen), Epogin (Chugai), Epomax (Eli Lilly), Eprex  
 (Janssen-Cilag), NeoRecormon or Recormon (Roche), and Procrit  
 (Ortho Biotech). Variations in the glycosylation pattern of EPO  
 distinguishes these products. Epogen, Epogin, Eprex and Procrit  
 are genetically known as epoetin alfa, NeoRecormon and Recormon as  
 epoetin beta and Epomax as epoetin omega.  
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.  
 CC -1- DATABASE: NAME=Red Systems' cytokine source book; EPO;  
 WWW="http://www.rndsystems.com/asp/g\_sitbuilder.asp?bodyid=197".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X02158; CA26095.1; -  
 DR EMBL; X02157; CA26094.1; -  
 DR EMBL; M11319; AAA52400.1; -  
 DR EMBL; AF053356; AAC78791.1; -

DR EMBL; AF202308; AAF23132.1; -  
 DR EMBL; AF202306; AAF23132.1; JOINED.  
 DR EMBL; AF202307; AAF23132.1; JOINED.  
 DR EMBL; AF202310; AAF23133.1; -  
 DR EMBL; AF202309; AAF23133.1; JOINED.  
 DR EMBL; AF202311; AAF21572.1; -  
 DR EMBL; AF202314; AAF23134.1; -  
 DR EMBL; AF202312; AAF23134.1; JOINED.  
 DR EMBL; AF202313; AAF23134.1; JOINED.  
 DR EMBL; S65458; AAD13964.1; -  
 DR PIR; A01855; Z0HU.  
 DR PIR; A25384; A25384.  
 DR PIR; A24744; A24744.  
 DR PIR; A22210; A22210.  
 DR PDB; 1EER; 01-OCT-99.  
 DR PDB; 1CN4; 11-AUG-99.  
 DR GlycoSuiteDB; P01588; -  
 DR GeneW; HGNC:3415; EPO.  
 DR MIM; 133170; -  
 DR InterPro; IPR001323; EPO\_TPO.  
 DR InterPro; IPR003013; Erythropo.  
 DR Pfam; PF00758; EPO\_TPO.1  
 DR PRINTS; PR00272; ERYTHROPTN.  
 DR PROSITE; PS00817; EPO\_TPO.1.  
 DR Erythrocyte maturation; Glycoprotein; Hormone; Signal; Pharmaceutical;  
 KW 3D-structure.  
 FT SIGNAL 1 27  
 FT CHAIN 28 193  
 FT PROPEP 190 193  
 FT DISULFD 34 188  
 FT DISULFD 56 60  
 FT CARBOHYD 51 51  
 FT CARBOHYD 65 65  
 FT CARBOHYD 110 110  
 FT CARBOHYD 153 153  
 FT CARBOHYD 131 132  
 FT VARIANT 149 149  
 FT VARIANT 149 149  
 FT CONFLICT 40 40  
 FT CONFLICT 85 85  
 FT CONFLICT 140 140  
 FT SEQUENCE 193 AA; 21306 MW; C91F0E4C26A52033 CRC64;  
 SO  
 Query Match 100.0%; Score 851; DB 1; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 7, 7e-77;  
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 APPLIDSVLELYLLLEAKENITTCGCHCSLNENITVPPTKYNFAMKREYGOQA 60  
 DB 28 APPLIDSVLELYLLLEAKENITTCGCHCSLNENITVPPTKYNFAMKREYGOQA 87  
 QY 61 VEWQGLALISEAVLRQALLVNSQWPEPLQIHVDKAVSGLSLTLLLRALGAQREAIS 120  
 DB 88 VEWQGLALISEAVLRQALLVNSQWPEPLQIHVDKAVSGLSLTLLLRALGAQREAIS 147  
 QY 121 PPDAAAPAPLRTTADTFRKLFVYSNFRGLKILKLYTGECRTGDR 166  
 DB 148 PPDAAAPAPLRTTADTFRKLFVYSNFRGLKILKLYTGECRTGDR 193  
 RESULT 2  
 ID EPO\_MACFA STANDARD: PRT; 192 AA.  
 AC P07865;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 16-OCT-1988 (Rel. 08, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Erythropoietin precursor.

GN EPO.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87055236; PubMed=2877922;  
RA Lin F.-K., Lin C.-H., Lai P.-H., Browne J.K., Eagle J.C., Smalling R.,  
Fox G.M., Chen K.K., Castro M., Suggs S.;  
RT "Monkey erythropoietin gene: cloning, expression and comparison with  
the human erythropoietin gene.";  
RL gene 44:201-209(1986).  
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE  
REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A  
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS  
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.  
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: M18189; AAA36841.1; -  
DR PIR: J00173; J00173.  
DR HSSP: P01588; ICN4.  
DR InterPro: IPR001323; EPO\_TPO.  
DR InterPro: IPR003013; Erythropo.  
DR Pfam: PF00758; EPO\_TPO; 1.  
DR PRINTS: PR00272; ERYTHROPTN.  
DR PROSITE: PS00817; EPO\_TPO; 1.  
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.  
FT SIGNAL 1 27  
FT CHAIN 1 27  
FT DISULFID 34 187  
FT DISULFID 56 60  
FT CARBOHYD 51 51  
FT CARBOHYD 65 65  
FT CARBOHYD 110 110  
FT CARBOHYD 152 152  
SQ SEQUENCE 192 AA; 21113 MW; E8A900F442AD4522 CRC64;  
  
Query Match 90.4%; Score 769.5; DB 1; Length 192;  
Best Local Similarity 91.6%; Pred. No. 8,3e-69;  
Matches 152; Conservative 7; Mismatches 6; Indels 1; Gaps 1;  
  
QY 1 APPRLICDSRVLEERYLLEAKEAENITTCGAHCESLSENIITVPTTKVNFYAMKREYGOOA 60  
DB 28 APPRLICDSRVLEERYLLEAKEAENITTCGAHCESLSENIITVPTTKVNFYAMKREYGOOA 87  
QY 61 VEVWQGLALSEAVLRSQALLVNSQWPEPLQHVDRKAVSGLSLITLLRALGAQKEAIS 120  
DB 88 VEVWQGLALSEAVLRSQALLVNSQWPEPLQHVDRKAVSGLSLITLLRALGAQ-PAIS 146  
QY 121 PPDAAAPLRTTTADTFKRLFRVYSNPLRGKILKLYTGACRGDR 166  
DB 147 LPDAAAPLRTTTADTFKRLFRVYSNPLRGKILKLYTGACRGDR 192  
  
RESULT 3  
EPO\_MACMU  
ID EPO\_MACMU STANDARD; PRT; 192 AA.  
AC Q28513;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Erythropoietin precursor.  
GN EPO.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93372347; PubMed=8364201;  
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,  
Czelusniak J., Goodman M., Bunn H.F.;  
RT "Erythropoietin structure-function relationships: high degree of  
sequence homology among mammals.";  
RL Blood 82:1507-1516(1993).  
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE  
REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A  
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS  
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.  
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: L10609; AAA36842.1; -  
DR HSSP: P01588; ICN4.  
DR InterPro: IPR001323; EPO\_TPO.  
DR InterPro: IPR003013; Erythropo.  
DR Pfam: PF00758; EPO\_TPO; 1.  
DR PRINTS: PR00272; ERYTHROPTN.  
DR PROSITE: PS00817; EPO\_TPO; 1.  
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.  
FT SIGNAL 1 27  
FT CHAIN 1 27  
FT DISULFID 34 187  
FT DISULFID 56 60  
FT CARBOHYD 51 51  
FT CARBOHYD 65 65  
FT CARBOHYD 110 110  
FT CARBOHYD 152 152  
SQ SEQUENCE 192 AA; 21081 MW; 275560A264628CD1 CRC64;  
  
Query Match 89.8%; Score 764.5; DB 1; Length 192;  
Best Local Similarity 90.4%; Pred. No. 2,6e-68;  
Matches 150; Conservative 9; Mismatches 6; Indels 1; Gaps 1;  
  
QY 1 APPRLICDSRVLEERYLLEAKEAENITTCGAHCESLSENIITVPTTKVNFYAMKREYGOOA 60  
DB 28 APPRLICDSRVLEERYLLEAKEAENITTCGAHCESLSENIITVPTTKVNFYAMKREYGOOA 87  
QY 61 VEVWQGLALSEAVLRSQALLVNSQWPEPLQHVDRKAVSGLSLITLLRALGAQKEAIS 120  
DB 88 VEVWQGLALSEAVLRSQALLVNSQWPEPLQHVDRKAVSGLSLITLLRALGAQ-PAIS 146  
QY 121 PPDAAAPLRTTTADTFKRLFRVYSNPLRGKILKLYTGACRGDR 166  
DB 147 LPDAAAPLRTTTADTFKRLFRVYSNPLRGKILKLYTGACRGDR 192  
  
RESULT 4  
EPO\_FEICA  
ID EPO\_FEICA STANDARD; PRT; 192 AA.  
AC P33708;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)

```
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Fells silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Fells.
OX NCBI_TaxID=9685;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RP Goodman R.E.; Bell R.G.;
  Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE OF 5-192 FROM N.A.
RA MEDLINE-93372347; PubMed-3364201;
  Wen D., Bolssel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
  Czelusniak J., Goodman M., Bunn H.F.,
  "Erythropoietin structure-function relationships: high degree of
  sequence homology among mammals.";
  Blood 82:1507-1516(1993).
RT
RL "FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
  REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
  PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS."
CC
CC -1- SUBCELLULAR LOCATION: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
  AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: U00685; AAA18282.1; -
DR EMBL: L10606; AAA10807.1; -
DR HSSP: P01588; 1CN4.
DR InterPro: IPR001323; EPO_TPO.
DR Pfam: PF00758; EPO_TPO; 1.
DR PRINTS: PS00272; ERYTHROPTN.
DR PROSITE: PS00817; EPO_TPO; 1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
CC
FT CHAIN 1 26
FT DISULFID 33 187
FT DISULFID 55 187
FT CARBOHYD 50 50
FT CARBOHYD 64 64
FT CARBOHYD 109 109
FT CONFLICT 44 109
FT CONFLICT 44 109
SQ SEQUENCE 192 AA; 20914 MW; 61C5EA0F5E937293 CRC64;
Query Match
Best Local Similarity 83.5%; Score 711; DB 1; Length 192;
Matches 139; Conservative 9; Mismatches 18; Indels 0; Gaps 0;
OY 1 APPRLICDSRVLEERYLLKEAKEENTTGAECSCSINENITVPDVKVNFYAKRMREVGQA 60
DB 27 APPRLICDSRVLEERYLLKEAKEENTTGAECSCSINENITVPDVKVNFYAKRMREVGQA 86
OY 61 VEWOGGLALISFAVLRGQALLVNSQPEPQLQHVDAVSGLSRSTLTLLRLAGOKRAIS 120
DB 87 VEWOGGLALISFAVLRGQALLVNSQPEPQLQHVDAVSGLSRSTLTLLRLAGOKRAIS 146
OY 121 PDAASAAPLRTITADTFRKILFRVSNFLRGKILKILYGEACRGCTGDR 166
DB 147 LPEATSAAPLRTITADTFRKILFRVSNFLRGKILKILYGEACRGCTGDR 192
RESULT 5
EPO_RAT ID EPO_RAT STANDARD; PRT; 192 AA.
AC P29676; P70504;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 16-OCT-2001 (Rel. 40, last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Kidney;
RP MEDLINE-93042015; PubMed-1420369;
  Nagao M., Suga H., Okano M., Masuda S., Narita H., Ikura K.,
  Sasaki R.;
  "Nucleotide sequence of rat erythropoietin.";
  Biochim. Biophys. Acta 1171:93-102(1992).
[2]
RN SEQUENCE OF 4-192 FROM N.A.
RA STRAIN=Sprague-Dawley; TISSUE=Kidney;
RP MEDLINE-93372347; PubMed-8364201;
  Wen D., Bolssel J.P.R., Tracy T.E., Mulcahy L.S., Czelusniak J.,
  Goodman M., Bunn H.F.,
  "Erythropoietin structure-function relationships: high degree of
  sequence homology among mammals.";
  Blood 82:1507-1516(1993).
RT
RL "FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
  REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
  PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS."
CC
CC -1- SUBCELLULAR LOCATION: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
  AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: D10763; BAA01592.1; -
DR EMBL: L10608; AAA1126.1; -
DR PUR: S28148; S28148.
DR HSSP: P01588; 1CN4.
DR InterPro: IPR001323; EPO_TPO.
DR Pfam: PF00758; EPO_TPO; 1.
DR PRINTS: PS00272; ERYTHROPTN.
DR PROSITE: PS00817; EPO_TPO; 1.
KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
CC
FT CHAIN 1 26
FT DISULFID 33 187
FT DISULFID 50 50
FT CARBOHYD 64 64
FT CARBOHYD 109 109
SQ SEQUENCE 192 AA; 21286 MW; 3EA632737E7D2443 CRC64;
Query Match
Best Local Similarity 82.5%; Score 706; DB 1; Length 192;
Matches 137; Conservative 13; Mismatches 16; Indels 0; Gaps 0;
OY 1 APPRLICDSRVLEERYLLKEAKEENTTGAECSCSINENITVPDVKVNFYAKRMREVGQA 60
DB 27 APPRLICDSRVLEERYLLKEAKEENTTGAECSCSINENITVPDVKVNFYAKRMREVGQA 86
OY 61 VEWOGGLALISFAVLRGQALLVNSQPEPQLQHVDAVSGLSRSTLTLLRLAGOKRAIS 120
```

Dbb 87 VEWMOGJLSTLEALQNALQANSQSPPESQLHIDKRAISLRSLTSLRLVIGAKNELMS 146

Oy 121 PPDAASAPLFTTTADTFRRKLFRVYSNFKRGLKYTGECRTGDGR 166  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Ddb 147 PPDATOAAPLRTLTADTECKLFVRYSNFLRGKTLKYTGECRRBDR 192

RESULT 6

| EPO_BOVIN ID  | STANDARD: | PRT: | 192 AA. |
|---|-----------|------|---------|
| AC PA4617;  |           |      |         |
| DT 01-FEB-1996 (Rel. 33, Created)   |           |      |         |
| DT 01-FEB-1996 (Rel. 33, Last sequence update)  |           |      |         |
| DT 16-OCT-2001 (Rel. 40, Last annotation update)  |           |      |         |
| DE Erythropoietin precursor.<br>EPO.  |           |      |         |
| GN Bos taurus (Bovine).   |           |      |         |
| OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;<br>Mammalia; Euarchontia; Cetartiodactyla; Ruminantia; Pecora; Bovidae;<br>Bovidae; Bos.   |           |      |         |
| OC NCBI_TaxID=9913;<br>[1]  |           |      |         |
| RN RP SEQUENCE FROM N.A.  |           |      |         |
| RC STRAIN=Boran; TISSUE=kidney;   |           |      |         |
| KX MEDLINE=96257233; PubMed=8666286;  |           |      |         |
| RA Suliman H.B., Majlwa P.A.O., Feldman B.F., Mertens B.,<br>Logan-Henfrey L.L.;<br>"Cloning of a cDNA encoding bovine erythropoietin and analysis of its<br>transcription in selected tissues.";<br>Gene 171:275-280(1996)."   |           |      |         |
| RL -1 FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE<br>REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A<br>PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.   |           |      |         |
| -1 SUBCELLULAR LOCATION: Secreted.  |           |      |         |
| CC CC TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS<br>AND BY LIVER OF FETAL OR NEONATAL MAMMALS.  |           |      |         |
| CC CC -1 SIMILARITY: BELONGS TO THE EPO / TPQ FAMILY.   |           |      |         |
| CC CC -----   |           |      |         |
| CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration<br>between the Swiss Institute of Bioinformatics and the EMBL Outstation -<br>the European Bioinformatics Institute. There are no restrictions on ways<br>to use by non-profit institutions as long as its content is in no way<br>modified and this statement is not removed. Usage by and for commercial<br>entities requires a license agreement (See http://www.isb-sib.ch/announce/<br>or send an email to license@isb-sib.ch). |           |      |         |
| CC CC -----   |           |      |         |
| DR EMBL; LA1354; BAB41268.1; -;   |           |      |         |
| DR EMBL; U04762; AAA86653.1; -;   |           |      |         |
| DR HSSP; P01588; ICN4.  |           |      |         |
| DR InterPro; IPR001323; EPO.TPQ.  |           |      |         |
| DR InterPro; IPR003013; Erythroptn.   |           |      |         |
| pfam: PF00758; EPO_TPQ: 1.  |           |      |         |
| DR PRINTS; PR00272; ERYTHROPOTN.  |           |      |         |
| DR PROSITE; PS00817; EPO_TPQ: 1.  |           |      |         |
| FW Erythrocyte maturation; Glycoprotein; Hormone; Signal.<br>POTENTIAL.   |           |      |         |
| FT SIGNAL 1 25  |           |      |         |
| FT CHAIN 32 192   |           |      |         |
| FT DISULFID 32 187  |           |      |         |
| FT FT BY SIMILARITY.  |           |      |         |
| FT DISULFID 54 58   |           |      |         |
| FT CARBOHYD 49 49   |           |      |         |
| FT CARBOHYD 63 63   |           |      |         |
| CABOHYD 108 108   |           |      |         |
| N-LINKED (GLCNAG. . .) (POTENTIAL)<br>N-LINKED (GLCNAG. . .) (POTENTIAL)<br>SEQUENCE 192 AA; 21075 MW; DBC419022EF7B483A CRC64;   |           |      |         |

Query Match 82.0%; Score 697.5; DB 1; Length 192;  
Best Local Similarity 83.2%; Pred. No. 1.le-61;  
Matches 139; Conservative 8; Mismatches 19; Indels 1; Gaps 1;

Oy 1 APPRLICDSRVLERLLTAKEAEENTTCGAECSCLENITNYDPDTVNVYANKRMVEVGQA 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Ddb 26 APARLICDSRVLERILLAREARENATTGCCAESCFSENIIVTPDKTVNFYANKRMVEVGQA 85

Oy 61 VEWMOGLALLSEAVLRGOALLVNVSQPMEPEPLQLAHVDAVSGLRSJTTLRLRALGAQKAIS 120

```

Db      86 LEWOGGLLLLSFALLRGCALLANASQPECEALRIHDKAVSGEIRSLTSLRLGAKOEKAI 145
Qy      121 PPDA-SAAPLRTITADPFRKLEFRVYSNFLRCKLTYTGACRTGDR 166
        ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      146 LPDAIPSAAPLRAPIYVDA LSKFRIYSNFLRCKLTYTGACRRBDR 192

RESULT 7
EPO_MOUSE
AC ID EPO_MOUSE STANDARD: PRT: 192 AA.
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A..
RX MEDLINE=87039105; Pubmed=3773894;
RA Shoemaker C.B., Miltsock L.D.;
RT "Murine erythropoietin gene: cloning, expression, and human gene
RT homology";
RL Mol. Cell. Biol. 6:849-858(1986).
RN [2]
RP SEQUENCE FROM N.A..
RX MEDLINE=87039104; Pubmed=3022133;
RA McDonald J.D., Lin F.-K., Goldwasser E.;
RT "Cloning, sequencing, and evolutionary analysis of the mouse
RT erythropoietin gene";
RL Mol. Cell. Biol. 6:842-848(1986).
RN [3]
RP SEQUENCE FROM N.A..
RC STRAIN=129/Sv;
RX MEDLINE=21138439; Pubmed=112939002;
RA Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,
RA Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
RA Miller W., Koop B.F.;
RT "Comparative analysis of the gene-dense ACHE/PPR2 region on human
RT chromosome 7q24 with the orthologous region on mouse chromosome 5.";
RL Nucleic Acids Res. 29:1352-1365(2001).
RN [4]
RP SEQUENCE OF 1-52 FROM N.A.
RC STRAIN=ICFW;
RX MEDLINE=98030528; Pubmed=9365246;
RA Christian S., Duprez V., Maouche L., Gisselbrecht S., Mayeux P.,
RA Lacombe C.;
RT "Abnormal erythropoietin (Epo) gene expression in the murine
RT erythroleukemia Iw32 cells results from a rearrangement between the
RT g-protein betaz2 subunit gene and the Epo gene.";
RL Oncogene 15:1995-1999(1997).
CC -I- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASSES.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -I- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; M12482; AAA37568.1; -
DR EMBL; M12930; AAA37570.1; -

```

```

DR EMBL; AF312033; AAK28825.1; -.
DR EMBL; U11971; CAA72707.1; -.
DR PIR; A24901; A24901.
DR PIR; A24902; A24902.
DR HSSP; P01588; 1CN4.
DR MGd; MGI:95407; EPO.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropn.
DR Pfam; PF00758; EPO_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO; 1.
KM Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 26
FT CHAIN 1 26
FT DISULFID 27 192
FT CARBOHYD 33 187
FT CARBOHYD 50 50
FT CARBOHYD 64 64
FT CARBOHYD 109 109
SEQUENCE 192 AA; 21365 MW; 65F94E214EDDEF2E CRC64;

Query Match
Best Local Similarity 81.6%; Score 694; DB 1; Length 192;
Matches 133; Conservative 14; Mismatches 19; Indels 0; Gaps 0;

QY 1 APPRLICDSRYLERYLLEAKENITTCGAHCSLNTNITVPDTKYNFYAKRMVGGQA 60
DB 27 APPRLICDSRYLERYLLEAKENITTCGAHCSLNTNITVPDTKYNFYAKRMVGGQA 86
QY 61 VEWOGIALISEAVLRGOALLVNSOPWEPLQHLVDKAVSGRSITTLIRALGAQKEAIS 120
DB 87 LEWOGIALISEAVLRGOALLVNSOPWEPLQHLVDKAVSGRSITTLIRALGAQKEAIS 146
QY 121 PDPAASAPLRTITADTFRKLFERYVSNFLRGKIKLYTGEACRGTGDR 166
DB 147 LPDPTTAPLRTITADTFRKLFERYVSNFLRGKIKLYTGEACRGTGDR 192

RESULT 8
EPO_SHEEP
ID EPO_SHEEP STANDARD; PRT: 194 AA.
AC P33709; Q28572;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor.
GN EPO.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RC MEDLINE=93351736; PubMed=8349021;
RA Fu P., Evans B., Lim G.B., Moritz K., Wintour M.E.;
RT "The sheep erythropoietin gene: molecular cloning and effect of
RT hemorrhage on plasma erythropoietin and renal/liver messenger RNA in
RT adult sheep.";
RL Mol. Cell. Endocrinol. 93:107-116(1993).
RN [2]
RP SEQUENCE OF 4-194 FROM N.A.
RC TISSUE-Kidney;
RC MEDLINE=93372347; PubMed=8364201;
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993)
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.

```

```

CC CC
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z24681; CAA80848.1; -.
DR EMBL; U10610; AAA31518.1; -.
DR HSSP; P01588; 1CN4.
DR InterPro; IPR001323; EPO_TPO.
DR InterPro; IPR003013; Erythropn.
DR Pfam; PF00758; EPO_TPO; 1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO; 1.
KM Erythrocyte maturation; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 27
FT CHAIN 1 27
FT DISULFID 28 194
FT CARBOHYD 34 189
FT CARBOHYD 56 60
FT CARBOHYD 51 51
FT CARBOHYD 65 65
FT CARBOHYD 110 110
FT CARBOHYD 16 16
FT CONFLICT 108 108
SEQUENCE 194 AA; 21335 MW; C025AAB0528131A9 CRC64;

Query Match
Best Local Similarity 82.0%; Score 690.5; DB 1; Length 194;
Matches 137; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

QY 1 APPRLICDSRYLERYLLEAKENITTCGAHCSLNTNITVPDTKYNFYAKRMVGGQA 60
DB 28 APPRLICDSRYLERYLLEAKENITTCGAHCSLNTNITVPDTKYNFYAKRMVGGQA 87
QY 61 VEWOGIALISEAVLRGOALLVNSOPWEPLQHLVDKAVSGRSITTLIRALGAQKEAIS 120
DB 88 LEWOGIALISEAVLRGOALLVNSOPWEPLQHLVDKAVSGRSITTLIRALGAQKEAIS 147
QY 121 PDPAASAPLRTITADTFRKLFERYVSNFLRGKIKLYTGEACRGTGDR 166
DB 148 LPDPTTAPLRTITADTFRKLFERYVSNFLRGKIKLYTGEACRGTGDR 194

RESULT 9
EPO_PIG
ID EPO_PIG STANDARD; PRT: 190 AA.
AC P49157;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor (Fragment).
GN EPO.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sub.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RC MEDLINE=93372347; PubMed=8364201;
RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA Czelusniak J., Goodman M., Bunn H.F.;
RT "Erythropoietin structure-function relationships: high degree of
RT sequence homology among mammals.";
RL Blood 82:1507-1516(1993)
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A

```

```

CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC      -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: L10607; AAA31029.1; -.
CC      DR      HSSP: P01588; 1CN4.
CC      DR      InterPro: IPR001323; EPO_TPO.
CC      DR      Pfam: PF00758; EPO_TPO; 1.
CC      DR      PROSITE: PS00817; EPO_TPO; 1.
CC      KM      Erythrocyte maturation; Glycoprotein; Hormone; Signal.
CC      FT      NON_TER 1
CC      FT      SIGNAL <1 22
CC      FT      CHAIN 23 190 ERYTHROPOIETIN.
CC      FT      DISULFID 29 185 BY SIMILARITY.
CC      FT      DISULFID 51 55 BY SIMILARITY.
CC      FT      CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT      CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT      CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT      CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      SQ      SEQUENCE 190 AA; 20888 MW; A75BD6CCE5077E2A CRC64;

Query Match      80.3%; Score 683; DB 1; Length 190;
Best Local Similarity 82.1%; Pred. No. 2.8e-60;
Matches 138; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

OY      1 APPRLICDSRVLEERYLLLEAKAEENITTCGAHCISLNEINIVPTTKVNFYAMKREMEYGOQA 60
DB      23 APPRLICDSRVLEERYLLLEAKAEENITTCGAHCISLNEINIVPTTKVNFYAMKREMEYGOQA 82
OY      61 VEWVQGLALISEAVNLGQALLVNSQPEWPELQIHDVKAVSGLSLTLRALGAQKEAIS 120
DB      83 MEVWQGLALISEAVNLGQALLVNSQPEWPELQIHDVKAVSGLSLTLRALGAQKEAIP 142
OY      121 PPDA--ASAPPLRTITADTFKRLFRVYSNPLRGKLTLYTGEACRRTDR 166
DB      143 LPDASSSATPLRTFAVDTLCKLFRNYSNPLRGKLTLYTGEACRRDR 190

RESULT 10
EPO_CANFA
ID      EPO_CANFA      STANDARD;      PRT;      175 AA.
AC      P33707;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Erythropoietin precursor (Fragment).
GN      EPO.
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-93372347; PubMed-8364201;
RA      Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
RA      Czelusniak J., Goodman M., Bunn H.F.;
RT      "Erythropoietin structure-function relationships: high degree of
RT      sequence homology among mammals.";
RL      Blood 82:1507-1516(1993).
CC      -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC      REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC      -1- SUBCELLULAR LOCATION: Secreted.

```

```

CC      -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC      -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: I13037; AAA30842.1; -.
CC      DR      HSSP: P01588; 1CN4.
CC      DR      InterPro: IPR001323; EPO_TPO.
CC      DR      Pfam: PF00758; EPO_TPO; 1.
CC      DR      PROSITE: PS00817; EPO_TPO; 1.
CC      KM      Erythrocyte maturation; Glycoprotein; Hormone; Signal.
CC      FT      NON_TER 1
CC      FT      SIGNAL <1 22
CC      FT      CHAIN 23 >175 ERYTHROPOIETIN.
CC      FT      DISULFID 29 >175 BY SIMILARITY.
CC      FT      DISULFID 51 55 BY SIMILARITY.
CC      FT      CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT      CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT      CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC      FT      NON_TER 175
CC      SQ      SEQUENCE 175 AA; 19193 MW; B504F8DBE676BF4 CRC64;

Query Match      75.0%; Score 638; DB 1; Length 175;
Best Local Similarity 81.0%; Pred. No. 7e-56;
Matches 124; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

OY      1 APPRLICDSRVLEERYLLLEAKAEENITTCGAHCISLNEINIVPTTKVNFYAMKREMEYGOQA 60
DB      23 APPRLICDSRVLEERYLLLEAKAEENITTCGAHCISLNEINIVPTTKVNFYAMKREMEYGOQA 82
OY      61 VEWVQGLALISEAVNLGQALLVNSQPEWPELQIHDVKAVSGLSLTLRALGAQKEAIS 120
DB      83 LEVWQGLALISEAVNLGQALLVNSQPEWPELQIHDVKAVSGLSLTLRALGAQKEAMS 142
OY      121 PPDAASAPPLRTITADTFKRLFRVYSNPLRGKLT 153
DB      143 LPDASAPPLRTITADTFKRLFRVYSNPLRGKLT 175

RESULT 11
TPO_CANFA
ID      TPO_CANFA      STANDARD;      PRT;      352 AA.
AC      P42705;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
DE      (c-MPL ligand) (ML) (Megakaryocyte growth and development factor)
GN      (MGDF).
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxID=9615;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 24-44.
RX      TISSUE-Kidney;
RA      MEDLINE-94291201; PubMed-8020099;
RA      Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.-S., Martin F.,
RA      Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
RA      Hsu R.-Y., Parker V.P., Sugas S., Skrine J.D., Merewether L.A.,
RA      Closson C., Hsu E., Hukom M.M., Hornkohl A., Choi E., Pangelinan M.,
RA      Sun Y., Mar Y., McNich J., Simonet L., Jacobsen F., Xie C.,
RA      Shutter J., Chute H., Basu R., Selander L., Trollinger D., Steu L.,
RA      Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
RA      Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,

```

[illegible]

|                       |                  |                   |       |             |
|-----------------------|------------------|-------------------|-------|-------------|
| Query Match           | 12.83;           | Score 109;        | DB 1; | Length 352; |
| Best Local Similarity | 24.0%;           | Pred. No. 0.0026; |       |             |
| Matches 40;           | Conservative 23; | Mismatches 23;    |       |             |

[illegible]

| RESULT 12 | TPQ_HUMAN   | STANDARD:                               | PM: | 353 AA. |
|-----------|---|---|-----|---------|
| ID        | TPQ_HUMAN   | P40325; Q13020; Q15790; Q15791; Q15792; |     |         |
| AC        | P40325; Q13020; Q15790; Q15791; Q15792;                             |   |     |         |
| DT        | 01-FEB-1995 (Rel. 31, Created)                                      |   |     |         |
| DT        | 01-FEB-1995 (Rel. 31, Last sequence update)                         |   |     |         |
| DT        | 15-JUN-2002 (Rel. 41, Last annotation update)                       |   |     |         |
| DE        | Thrombopoietin precursor (Megakaryocyte colony stimulating factor)  |   |     |         |
| DE        | (Meloproliferative leukemia oncogene ligand) (c-mpl ligand)         |   |     |         |
| DE        | (Myelopoietic growth and development factor) (MDFP).                |   |     |         |
| GN        | THPO.   |   |     |         |
| OS        | Homo sapiens (Human).   |   |     |         |
| OC        | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |   |     |         |
| OC        | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.           |   |     |         |
| OX        | NCBI_Taxid:9606;  |   |     |         |
| RN        | [1]   |   |     |         |
| RP        | SEQUENCE FROM N.A. (ISOFORM 1).                                     |   |     |         |
| RC        | TISSUE: Fetal liver;  |   |     |         |
| XX        | MEDLINE:94261202; PubMed:8202154;                                   |   |     |         |
| RA        | De Sauvage F.J., Hass P.E., Spencer S.D., Malloy B.E., Gurney A.L., |   |     |         |
| RA        | Spencer S.A., Darbonne W.C., Henzel W.J., Wong S.C., Kuang W.-J.,   |   |     |         |
| RA        | Oles K.J., Hultgren B., Solberg L.A. Jr., Goeddel D.V., Eaton D.L., |   |     |         |
| RT        | Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mp  |   |     |         |
| TI        | ligand";  |   |     |         |

RL Nature 369:533-538(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Fetal liver;  
 RX MEDLINE=94291201; PubMed=8020099;  
 RA Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,  
 RA Chang M.-S., Samel B.B., Nichol J.L., Swift S., Johnson M.J.,  
 RA Hou R.-T., Parker V.P., Suggs S., Skrine J.D., Merwether I.A.,  
 RA Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangelinan M.,  
 RA Sun Y., Mar V., McNich J., SImonet L., Jacobsen F., Xie C.,  
 RA Shutter J., Chute H., Basu R., Selander L., Trollinger D., Sien L.,  
 RA Padilla A., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,  
 RA Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,  
 RA Pacifici R., Ponting I., Saris C., Wen D., Yung I.P., Lin H.,  
 RA Bosselman R.A.;  
 RT "Identification and cloning of a megakaryocyte growth and development  
 RT factor that is a ligand for the cytokine receptor Mpl."  
 RL Cell 77:1117-1124(1994).  
 RN [3]

RX SEQUENCE FROM N.A. (ISOFORM 1).  
 RP MEDLINE=95108091; Pubmed=7809166;  
 RA Foster D.C., Sprecher C.A., Grant F.J., Kramer J.M., Kuipjer J.L.,  
 RA Molloy R.D., Whitmore T.E., Heipel M.D., Bell L.A.N., Ching A.F.,  
 RA Humane T.V., Hart G., O'Hara P.J., Lok S.;  
 RT "Human tyrosinase: gene structure, cDNA sequence, expression, and  
 RT chromosomal localization."  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:13023-13027(1994).  
 [4]

RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA MEDLINE=95010765; PubMed=7926003;  
 RA Motuma Y., Akahori H., Seki N., Hori T.-A., Ogami K., Kawamura K.,  
 RT "Molecular cloning and chromosomal localization of the human  
 RT thrombopoietin gene." *FEBS Lett.* 353:57-61(1994).  
 RL [5]  
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RP MEDLINE=95152076; PubMed=7843919;  
 RA Gunney A.L., Kuang W.-J., Xie M.-H., Malloy B.E., Baton D.L.,  
 RT "Genomic structure, chromosomal localization, and conserved  
 RT alternative splice forms of thrombopoietin." *Blood* 85:981-988(1995).

RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Liver;  
RX MEDLINE-96015174; PubMed-6537317;  
RA Kato T., Osami K., Shimada Y., Iwamatsu A., Sohma Y., Akahori H.,  
RA Kikuchi K., Kokubo A., Kudo Y., Meeda E., Kobayashi K., Ohashi H.,  
RA Ozawa T., Inoue H., Kawamura K., Miyazaki H.;  
RT "Purification and characterization of thrombopoietin.";  
RL J. Biochem. 118:229-236(1995).  
RN [7].  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Placenta;  
RX MEDLINE-95122483; PubMed-7822271;  
RA Chang M., McIninch J., Basu R., Shutter J., Hsu R., Perkins C., Mar V.,  
RA Suggs S., Welcher A., Li L., Lu H., Bartley T., Hunt P., Martin F.,  
RA Samuel B., Bogenberger J.;  
RT "Cloning and characterization of the human megakaryocyte growth and  
RT development factor (MGDF) gene.";  
RL J. Biol. Chem. 270:511-514(1995).  
RN [8].

RA SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).  
 RT Im S.H., Lee W.S., Chung K.H.;  
 PT Cloning and sequencing of human thrombopoietin."  
 CC Submitted (May-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1 FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION  
 CC OF MEGAKARYOCYTES FROM THEIR COMMITTED PRECURSOR  
 CC CELLS. INDUCTION OF MEGAKARYOCYTES FROM THEIR COMMITTED PRECURSOR  
 CC MAY BE THE FACTOR AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT  
 CC -1 SURCELS THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS  
 CC -1 ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/TPO-2 AND



```
CC 3/TRUNCATED; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -1- DATABASE: NAME-RED systems' cytokine source book: TPO;
CC WWW="http://www.rndsystems.com/asp/9/sitebuilder.asp?bodyid=225".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L33410; AAA59857.1; -
DR EMBL: U11025; AAA50553.1; -
DR EMBL: L36051; AAC37568.1; -
DR EMBL: L36052; AAC37566.1; -
DR EMBL: D32046; BAA06807.1; -
DR EMBL: S76771; AAB33390.1; -
DR EMBL: D32047; BAA21830.1; -
DR EMBL: U59493; AAB03392.1; -
DR EMBL: U59494; AAB03393.1; -
DR EMBL: U59495; AAB03394.1; -
DR EMBL: U11071; AAA74083.1; -
DR PIR: S45331; S45331.
DR Genew: HGNC:11795; THPO.
DR MIM: 600044; -.
DR InterPro: IPR001323; EPO_TPO.
DR InterPro: IPR003978; thrombopoietin.
DR Pfam: PF00758; EPO_TPO; 1.
DR PRINTS: PR01485; THROMBOPTN.
DR PROSITE: PS00817; EPO_TPO; 1.
KM Cytokine; Glycoprotein; Hormone; Signal; Alternative splicing;
KM Polymorphism.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 353 THROMBOPOIETIN.
FT DISULFD 28 172 POTENTIAL.
FT DISULFD 50 106 POTENTIAL.
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 133 136 MISSING (IN ISOFORM 2).
FT VARSPLIC 160 198 MISSING (IN ISOFORM 3).
FT VARIANT 14 14 L-> P (IN DBSNP:1042346).
FT VARIANT 14 14 /FTID=VAR.011795.
FT VARIANT 116 116 G-> E (IN DBSNP:1126665).
FT VARIANT 116 116 /FTID=VAR.011796.
FT CONFLICT 46 46 R-> K (IN REF. 8).
FT CONFLICT 76 76 M-> MSQ (IN REF. 7).
FT CONFLICT 113 113 Q-> E (IN REF. 2).
FT CONFLICT 131 131 T-> P (IN REF. 7).
FT CONFLICT 277 277 G-> E (IN REF. 8; AAB03393/AAB03394).
FT CONFLICT 346 346 S-> C (IN REF. 8; AAB03393/AAB03394).
SQ SEQUENCE 353 AA: 37822 MW: 49825.526 CRC64;

Query Match 10.5%; Score 89; DB 1; Length 353;
Best Local Similarity 26.3%; Pred. No. 0.24;
Matches 41; Conservative 20; Mismatches 75; Indels 20; Gaps 5;
```

```
DB 133 -LPPQG-----RTAHNDPMNIFLSFOHLNGKRA 161
|||
RESULT 13
CH60_BUCMP STANDARD; PRT; 548 AA.
ID C51832; BUCMP STANDARD; PRT; 548 AA.
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa chaperonin (Protein Cpn60) (groEL protein) (Symbionin).
GN GRO1 OR MODA OR GROEL OR SYML.
OS Buchnera aphidicola (subsp. Myzus persicae) (Myzus persicae primary
endosymbiont).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=98795;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98080423; PubMed=9420234;
RA Hogenhout S.A., van der Wilk F., Verbeek M., Goldbach R.W.,
RA van den Heuvel J.F.J.M.;
RA "Potato leafroll virus binds to the equatorial domain of the aphid
RT endosymbiotic GROEL homolog.";
RL J. Virol. 72:358-365(1998)
CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF003957; AAC04237.1; -.
DR HSSP: P06139; LJON.
DR InterPro: IPR001844; Chaperin_Cpn60.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR Pfam: PF00118; cpn60_TCP1.1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.
KM Chaperone; ATP-binding.
SQ SEQUENCE 548 AA: 57883 MW: AD61917965E4184A CRC64;

Query Match 9.6%; Score 82; DB 1; Length 548;
Best Local Similarity 24.2%; Pred. No. 2.1;
Matches 43; Conservative 33; Mismatches 70; Indels 32; Gaps 8;

QY 10 RYLERLLLEA-KEAENITTCGAHCNSINENITV--PDKVNYAKRRE-VGQAV-EY 63
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 118 RQIDRAVISAVEELKRLSPVCSDSKAITGYGTISANADERVGALEAEKEKGNQGVTV 177
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 WQGLALLSE-AVLRGQALLVNSSQPW-----EPLQLDVKAIVSGRLSTLTLL 109
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 EESTGQNLNLEVYKQGQFPGYLSPYFIKKPENGIVLENPYLVADKISNVRMLPLI 237
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 110 RALGAQKEAISPPDAASAAPLRTTAD-TFRKLFVYNSFLRGKILKLYGCAECRTGDR 166
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 238 ESIV-----AKSGKPLIISBDELGAEALATLVNMSRGIVAAVAKAFGFDR 284
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
CH60_BUCAP STANDARD; PRT; 551 AA.
ID C59177; BUCAP STANDARD; PRT; 551 AA.
AC 059177;
DT 01-NOV-1997 (Rel. 35, Created)
```



```
FT CHAIN 1 115 CAPSID PROTEIN C (POTENTIAL).
FT CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
FT CHAIN 192 383 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
FT CHAIN 384 733 NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
FT CHAIN 734 1010 NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
FT CHAIN 1011 1619 PROTEASE/HELICASE NS3 (POTENTIAL).
FT CHAIN 1620 1866 NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
FT CHAIN 1867 2017 NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
FT CHAIN 2018 3033 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
FT TRANSMEM 347 369 POTENTIAL.
FT ACT_SITE 1087 1087 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1111 1111 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1169 1169 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1234 1241 ATP (POTENTIAL).
FT SITE 1320 1323 DECH BOX.
FT CARBOHYD 186 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 558 558 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1091 1091 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2038 2038 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2359 2359 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2811 2811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 3033 AA; 330177 MW; 1A173E7E381FD1A CRC64;
```

Query Match 9.5%; Score 80.5; DB 1; Length 3033;  
Best Local Similarity 26.6%; Pred. No. 24;

Matches 41; Conservative 22; Mismatches 64; Indels 27; Gaps 7;

```
QY 5 LICDSHFLERYLLLEAKENITFGCAE--HCSINENITV-PDTKVNIFYAMKRM----- 54
Db 1659 IMTSSWVLGAGVLAANVAICLATGICISITGRHLINDRVYVAPDKELIYAFDEMECCASK 1718
QY 55 ----EYGOQAVEVMQG--LALLSEAVLRGQALLVNSSQPEPLQ----LHYDKAVSGLRS 104
Db 1719 AALIEGQRMWAEMLKSKIGLQQAATROADIQPAIQSSWPKLEQFWAKHMMNFISGIQY 1778
QY 105 LTTLLRALG----AQKEAISPPDAASAAPLRTIT 134
Db 1779 LAGLSTLPGNPVAVASMAFS---AALTSPLPSTT 1809
```

Search completed: January 2, 2003, 15:11:53  
Job time : 10.5257 secs

**THIS PAGE BLANK (USPTO)**



OY 121 PPDA-SAAPLRTTADTFKRLFRVYSNPLRGKLTLYTGACRGR 166  
 149 PPEAASSAAPLRTVAADTLCKLFRVYSNPLRGKLTLYTGACRGR 195

## RESULT 2

OY 09GK2 PRELIMINARY; PRT; 195 AA.  
 AC 09GK2;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Erythropoietin.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA MEDLINE=21290682; PubMed=11396976;  
 RA Viala A., Ma D., Margalich M., Hobart P.;  
 RT Rabbit EPO gene and cDNA: Expression of Rabbit EPO after  
 RT Intramuscular Injection of PDNA.  
 RL Biochem. Biophys. Res. Commun. 284:823-827(2001).  
 DR HSP; P01588; ICN4.  
 DR InterPro: IPR001323; EPO\_TPO.  
 DR InterPro: IPR001323; Erythropo.  
 DR Pfam: PF00758; EPO\_TPO; 1.  
 DR PRINTS: PR00272; ERYTHROPTN.  
 DR PROSITE: PS00817; EPO\_TPO; 1.  
 SO SEQUENCE 195 AA; 21025 MW; 1F1DC7F403A03BC CRC64;

Query Match 80.6%; Score 685.5; DB 6; Length 195;  
 Best Local Similarity 81.4%; Pred. No. 8.6e-61;  
 Matches 136; Conservative 12; Mismatches 18; Indels 1; Gaps 1;

OY 1 APRLLICDSVRLRYLLEAKENITTCGAHCNSLNTITVPDTRVNFYAMKRMVGOQA 60  
 DB 29 APARLLICDSVRLRYLLEAKENITTCGAHCNSLNTITVPDTRVNFYAMKRMVGOQA 88  
 OY 61 VEWOGIALISEAVLRGQALLVNSQPEWPLQHLVDKAVSGLSLTTLRALGAQKEAIS 120  
 DB 89 VEWOGIALISEAVLRGQALLVNSQPEWPLQHLVDKAVSGLSLTTLRALGAQKEAIS 148  
 OY 121 PPDA-SAAPLRTTADTFKRLFRVYSNPLRGKLTLYTGACRGR 166  
 DB 149 PPEAASSAAPLRTVAADTLCKLFRVYSNPLRGKLTLYTGACRGR 195

## RESULT 3

OY 09MYM8 PRELIMINARY; PRT; 194 AA.  
 AC 09MYM8;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Erythropoietin precursor.  
 GN EPO.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA STRAIN=NORWEGIAN LANDRACE; TISSUE=KIDNEY;  
 RA David B., Harditz I.;  
 RT "The porcine erythropoietin gene: cDNA and genomic sequences and  
 RT expression analyses."  
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249745; CAB96416.1; -  
 DR EMBL; AJ249746; CAB96417.1; -

DR HSP; P01588; ICN4.  
 DR InterPro: IPR001323; EPO\_TPO.  
 DR InterPro: IPR001323; Erythropo.  
 DR Pfam: PF00758; EPO\_TPO; 1.  
 DR PRINTS: PR00272; ERYTHROPTN.  
 DR PROSITE: PS00817; EPO\_TPO; 1.  
 KW Signal.  
 FT SIGNAL 1 26  
 FT CHAIN 27 194  
 SO SEQUENCE 194 AA; 21303 MW; 77881A6F620E1C CRC64;

Query Match 80.3%; Score 683; DB 6; Length 194;  
 Best Local Similarity 82.1%; Pred. No. 1.6e-60;  
 Matches 138; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

OY 1 APRLLICDSVRLRYLLEAKENITTCGAHCNSLNTITVPDTRVNFYAMKRMVGOQA 60  
 DB 27 APARLLICDSVRLRYLLEAKENITTCGAHCNSLNTITVPDTRVNFYAMKRMVGOQA 86  
 OY 61 VEWOGIALISEAVLRGQALLVNSQPEWPLQHLVDKAVSGLSLTTLRALGAQKEAIS 120  
 DB 87 VEWOGIALISEAVLRGQALLVNSQPEWPLQHLVDKAVSGLSLTTLRALGAQKEAIS 146  
 OY 121 PPDA-SAAPLRTTADTFKRLFRVYSNPLRGKLTLYTGACRGR 166  
 DB 147 LPDAPSSAPLRTVAADTLCKLFRVYSNPLRGKLTLYTGACRGR 194

## RESULT 4

OY 09QV40 PRELIMINARY; PRT; 50 AA.  
 AC 09QV40;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Erythropoietin (Fragment).  
 OS Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE

RA MEDLINE=94115047; PubMed=7764337;  
 RA Okano M., Suga H., Masuda S., Nagao M., Narita H., Ikura K.,  
 RA Sasaki K.;  
 RL Biosci. Biotechnol. Biochem. 57:1882-1885(1993).  
 DR HSP; P01588; IER.  
 DR InterPro: IPR001323; EPO\_TPO.  
 DR InterPro: IPR001323; Erythropo.  
 DR Pfam: PF00758; EPO\_TPO; 1.  
 DR PRINTS: PR00272; ERYTHROPTN.  
 SO SEQUENCE 50 AA; 5587 MW; 70B44A8BFE016034 CRC64;

Query Match 22.1%; Score 188; DB 11; Length 50;  
 Best Local Similarity 78.0%; Pred. No. 1.2e-11;  
 Matches 39; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 APRLLICDSVRLRYLLEAKENITTCGAHCNSLNTITVPDTRVNFYAMKRMVGOQA 50  
 DB 1 APRLLICDSVRLRYLLEAKENITTCGAHCNSLNTITVPDTRVNFYAMKRMVGOQA 50

## RESULT 5

OY 08ZDC8 PRELIMINARY; PRT; 323 AA.  
 AC 08ZDC8;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Ribonucleoside-diphosphate reductase 2 beta chain (EC 1.17.4.1).  
 GN NRDF OR YPO2648.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

RA STRAIN=NORWEGIAN LANDRACE; TISSUE=KIDNEY;  
 RA David B., Harditz I.;  
 RT "The porcine erythropoietin gene: cDNA and genomic sequences and  
 RT expression analyses."  
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249745; CAB96416.1; -  
 DR EMBL; AJ249746; CAB96417.1; -







Best Local Similarity 23.8%; Pred. No. 3.4;  
Matches 46; Conservative 34; Mismatches 56; Indels 57; Gaps 10;

QY 6 ICDSRVLEKYLEAK-EAENITTCGAHCSLNENTIV--VDITKVNRYAMKRMV-GQOAV-EY 63  
 DB 32 LCDADTTKVOILTALNAESTLVDNA--ASLDALITMLPNSSE---AVSQVLLGSDGISG 85  
 QY 64 WQGLALLSE-AVLRGQALLVNSSQPM-----WE-----PQLQHVDAKAVSG----- 101  
 DB 86 W--VAOLSOAAV---VIDMSSSDPERSRLAILLAWEVDYDAPYGGVKAQNGTSLI 140  
 QY 102 -----LRSLTTLRALGAQKEAISPPDASAAPLRTITADTFKRLFRVYSNFLRGKLT 153  
 DB 141 LIGGEDRVLKSCYCTALAAANGEQILFVGPRAGSGHAA-----KALNNIYSAIG 186  
 QY 154 KLYTGEACRTGDR 166  
 DB 187 LATITELHVAQR 199

## RESULT 11

Q93T48 PRELIMINARY; PRT; 548 AA.  
 AC Q93T48;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE groEL-related molecular chaperonin SymL.  
 GN Buchnera aphidicola (subsp. Myzus persicae) (Myzus persicae primary endosymbiont).  
 OS Buchnera; Proteobacteria; gamma subdivision; Buchnera.  
 OC NCBI\_TaxID=98795;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cui X., Mu Y., Lin L.;  
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF367248; AAK52957.1; -  
 DR InterPro: IPR001844; Chaperin\_Cpn60.  
 DR InterPro: IPR002423; Cpn60\_TCP-1.  
 DR Pfam: PF00118; Cpn60\_TCP1.1.  
 DR PROSITE: PS00296; CHAPERONINS\_CPN60; UNKNOWN\_1.  
 KW ATP-binding; Chaperone.  
 SQ SEQUENCE 548 AA; 57921 MW; 47CC05E9BF3CA7C CRC64;

Query Match 9.8%; Score 83; DB 2; Length 548;

Best Local Similarity 24.2%; Pred. No. 8.4;  
Matches 43; Conservative 33; Mismatches 70; Indels 32; Gaps 8;

QY 10 RVLEKYLLEA-KEAENITTCGAHCSLNENTIV--PDKVNFYAMKRMV-GQOAV-EY 63  
 DB 118 RGLDKAVISAVELKMLSVPCSDSKAITQVGTISANADEKVGALLAEAMEKVGNDGVITV 177  
 QY 64 WQGLALLSE-AVLRGQALLVNSSQPM-----EPLQHVDAKAVSGRLSTLTL 109  
 DB 178 EESTGIONLEEVKGMQFDRGLSPFYFKPKPGVYELNPYILMADKISNVREMLPIL 237  
 QY 110 RALGAQKEAISPPDASAAPLRTITAD-TFRKLFVYSNFLRGKLTLYGECRTGDR 166  
 DB 238 ESV-----AKSGRPLIISDELEGEALATLVNMSMGIVAAVKAAPGFDR 284

## RESULT 12

Q93N35 PRELIMINARY; PRT; 548 AA.  
 AC Q93N35;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE groEL-related molecular chaperonin SymL.  
 GN Buchnera aphidicola (subsp. Rhopalosiphum maidis).  
 OS Buchnera; Proteobacteria; gamma subdivision; Buchnera.

NCBI\_TaxID=118109;

QY 6 ICDSRVLEKYLEAK-EAENITTCGAHCSLNENTIV--VDITKVNRYAMKRMV-GQOAV-EY 63  
 DB 32 LCDADTTKVOILTALNAESTLVDNA--ASLDALITMLPNSSE---AVSQVLLGSDGISG 85  
 QY 64 WQGLALLSE-AVLRGQALLVNSSQPM-----WE-----PQLQHVDAKAVSG----- 101  
 DB 86 W--VAOLSOAAV---VIDMSSSDPERSRLAILLAWEVDYDAPYGGVKAQNGTSLI 140  
 QY 102 -----LRSLTTLRALGAQKEAISPPDASAAPLRTITADTFKRLFRVYSNFLRGKLT 153  
 DB 141 LIGGEDRVLKSCYCTALAAANGEQILFVGPRAGSGHAA-----KALNNIYSAIG 186  
 QY 154 KLYTGEACRTGDR 166  
 DB 187 LATITELHVAQR 199

Query Match 9.8%; Score 83; DB 2; Length 548;  
Best Local Similarity 24.2%; Pred. No. 8.4;  
Matches 43; Conservative 33; Mismatches 70; Indels 32; Gaps 8;

QY 10 RVLEKYLLEA-KEAENITTCGAHCSLNENTIV--PDKVNFYAMKRMV-GQOAV-EY 63  
 DB 118 RGLDKAVISAVELKMLSVPCSDSKAITQVGTISANADEKVGALLAEAMEKVGNDGVITV 177  
 QY 64 WQGLALLSE-AVLRGQALLVNSSQPM-----EPLQHVDAKAVSGRLSTLTL 109  
 DB 178 EESTGIONLEEVKGMQFDRGLSPFYFKPKPGVYELNPYILMADKISNVREMLPIL 237  
 QY 110 RALGAQKEAISPPDASAAPLRTITAD-TFRKLFVYSNFLRGKLTLYGECRTGDR 166  
 DB 238 ESV-----AKSGRPLIISDELEGEALATLVNMSMGIVAAVKAAPGFDR 284

## RESULT 13

Q9DHD6 PRELIMINARY; PRT; 3033 AA.  
 AC Q9DHD6;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)].  
 GN Hepatitis C virus type 2b.  
 OS Hepatitis C virus type 2b.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirinae.  
 OC NCBI\_TaxID=31650;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE:21296595; PubMed:11402859;  
 RA WuraKant K., Abe M., Kageyama T., Kamoshita N., Nomoto A.;  
 RT "Down-regulation of translation driven by hepatitis C virus internal ribosomal entry site by the 3' untranslated region of RNA.";  
 RL Arch. 146:729-741(2001).  
 CC -1- Similarity: NO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
 DR EMBL: AB030907; BAB08107.1; -  
 DR HSP: P27958; IAIIV.  
 DR InterPro: IPR000345; Cyto\_heme\_bind.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR002522; HCV\_capsid.  
 DR InterPro: IPR002521; HCV\_core.  
 DR InterPro: IPR002519; HCV\_env.  
 DR InterPro: IPR002511; HCV\_NS1.  
 DR InterPro: IPR002518; HCV\_NS2.  
 DR InterPro: IPR004109; HCV\_NS3.  
 DR InterPro: IPR000745; HCV\_NS4a.  
 DR InterPro: IPR001490; HCV\_NS4b.  
 DR InterPro: IPR002868; HCV\_NS4c.  
 DR InterPro: IPR002166; HCV\_NS4d.  
 DR InterPro: IPR01543; HCV\_core.  
 DR Pfam: PF01542; HCV\_core.1.  
 DR Pfam: PF01539; HCV\_env.1.  
 DR Pfam: PF01560; HCV\_NS1.1.  
 DR Pfam: PF01538; HCV\_NS2.1.

```

DR Pfam: PF01006; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00998; HCV NS5b; 1.
DR Pfam: PF0186062; HCV NS1; 1.
DR SMART: SM00487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN_1.
KW Coat protein; Envelope protein; glycoprotein; Nonstructural protein;
FT Polyprotein; RNA-directed RNA polymerase; Transmembrane.
FT CHAIN 1 191 CORE PROTEIN.
FT CHAIN 192 383 E1 PROTEIN.
FT CHAIN 384 750 E2 PROTEIN.
FT CHAIN 751 913 E3 PROTEIN.
FT CHAIN 914 1030 NS2 PROTEIN.
FT CHAIN 1031 1061 NS3 PROTEINASE/HELICASE.
FT CHAIN 1062 1115 NS4 PROTEIN.
FT CHAIN 1116 1176 NS4B PROTEIN.
FT CHAIN 1177 1242 NS5A PROTEIN.
FT CHAIN 1243 3033 NS5B RNA-DEPENDENT RNA POLYMERASE.
SQ SEQUENCE 3033 AA; 329991 MW; 6B183FED090872B4 CRC64;

Query Match
Best Local Similarity 27.0%; Score 82.5; DB 12; Length 3033;
Matches 40; Conservative 21; Mismatches 60; Indels 27; Gaps 7;

QY 11 VLERYLEAKAEENITTCGAF--HCSINENITV-PDTKVFYAMKRM-----EV 56
Db 1665 VLAGGVLAAYVAYGATGCTSTIGRIHLNDQVYVAPKREITLYAEFEMECCASKALIEE 1724

QY 57 GQQAWEVWQ--LALSEAVIRGOALLVNSSQPEWELQ---LHVDKAVSGLSLTTLR 110
Db 1725 GQRMAEMIKKLTLLDQAKRKQADIDPAMQSSMPKIEQFAMHNMNFTSGIYLAGLST 1784

QY 111 ALG----AQKALISPPDASAPLRTTT 134
Db 1785 LPGNFAYASMAFAS--AALTSPLEPST 1809

RESULT 14
O9FRK91 PRELIMINARY; PRT; 815 AA.
AC O9FRK91;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 19, Last sequence update)
DE Seed maturation protein PM38 protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN 111
RF SEQUENCE FROM N.A.
RX STRAIN=COLUMBIA;
RK MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Nabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Physically assigned P1 and TAC clones.";
RL DNA Res. 5:203-216(1998).
EMBL: AB012244; BAB09119.1;
DR InterPro: IPR002086; Aldehyde_dehydr.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS00170; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00172; BRCT; 1.
SQ SEQUENCE 815 AA; 91534 MW; 4B1B602057D46B60 CRC64;

```

```

Query Match
Best Local Similarity 9.6%; Score 82; DB 10; Length 815;
Matches 42; Conservative 29; Mismatches 52; Indels 70; Gaps 12;

QY 21 EAENITTCGA-----EHCINENITVPDTKVFYAMKRMEEVG-----QQAWEVW 64
Db 328 EKDGLVNCASFSCIDLGKGRNEYC-IMQLVTPDSMLNMT-FRKGKVGDDPAEERLEEW 385

QY 65 OGALLSEAVIRGOALLV-----NSSQPEW-----PLQ---LHVDKAVSGL 102
Db 386 ED---EBAIKFALFELIAGNEFEFEWEREKTKQKPKHFFITMDGIEVSGALGL 441

QY 103 RSL-----TTLRALGAQ-----KEAISPDASAPLRTITADTFKRL 141
Db 442 RQLGTSAMCKLDSFVAFNFTKVLGQGEIYNVALMELGIDPD---LPMGLMTDIIHKRC 497

QY 142 FRYYSNPLGKLR 154
Db 498 EYVLEFVE-KVK 509

RESULT 15
O93N34 PRELIMINARY; PRT; 548 AA.
AC O93N34;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
GN SYNM.
OS Buchnera aphidicola (subsp. Rhopalosiphum padi).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OC NCBI_TaxID=98793;
RN 111
RF SEQUENCE FROM N.A.
RX W. X., Cui X., Lin L.;
RT "Rhopalosiphum padi endosymbiont GroEL-related molecular chaperonin
RT SyNM (synt) gene.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF387864; AAK62971.1;
DR InterPro: IPR001844; Chaperonin_Cpn60.
DR InterPro: IPR002423; Cpn60_TCP-1.
DR Pfam: PF00118; Cpn60_TCP1; 1.
DR PROSITE: PS00296; CHAPERONINS_CPN60; UNKNOWN_1.
KW ATP-binding; Chaperone.
SQ SEQUENCE 548 AA; 57752 MW; C0296E8B6EF2794 CRC64;

Query Match
Best Local Similarity 9.4%; Score 80; DB 2; Length 548;
Matches 43; Conservative 32; Mismatches 71; Indels 32; Gaps 8;

QY 10 RVLERYLLEA-KEAENITTCGAHCSINENITV---PDTKVFYAMKRMV-VGQAV-EV 63
Db 118 RGLDKAVISAVELKMLSVPCSDSKAITQGTISANADEKVCALIAEMEKVGNCGVITV 177

QY 64 WGLIALLSE-AVLRGOALLVNSSQPEW-----EPLQVHVDKAVSGLSLTTLR 109
Db 178 EESTGTQNELEVKKQGFQDGLVSPYFINKPETGVLEENPYITLMAADKISNVREMLPIL 237

QY 110 RALGAQKEAISPDASAPLRTITAD--TFKRLFVYSNFTLGKIKLITYGACRTGDR 166
Db 238 ESV-----AKSGKPLIISDEDEGEALATLVNSTRGIVKVAAYKAPFGDR 284

Search completed: January 2, 2003, 15:12:56
Job time : 28.5801 secs

```